REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Information Technology

OPTIMIZATION SMART CLASSROOM USING FEATURE SELECTION IN ACO ALGORITHM

Dhuha Abd Ulameer Abd Ali ABD ALI

Master's Thesis

Supervisor

Prof. Dr. Hasan Hüseyin BALIK

Istanbul, 2022

OPTIMIZATION SMART CLASSROOM USING FEATURE SELECTION IN ACO ALGORITHM

Dhuha Abdulameer Abd Ali ABD ALI

Information Technologies

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis/dissertation titled OPTIMIZATION SMART CLASSROOM USING FEATURE SELECTION IN ACO ALGORITHM prepared by DHUHA ABD ULAMEER ABD ALI ABD ALI and submitted 16/12/2022 has been **accepted unanimously** for the degree of Master of Science in Information Technologies.

____

Prof. Dr. Hasan Hüseyin BALIK

The Supervisor

Thesis Défense Committee Members:

Prof. Dr. Hasan Hüseyin BALIK	Department of Computer Engineering Yildiz Technique University	____________________
Asst. Prof. Dr. Oguz ATA	Department Of Software Engineering Altinbas University	____________________
Asst.Prof.Dr. Aytug BOYACI	Department of Computer Engineering National Defense University	____________________

I hereby declare that this thesis meets all format and submission requirements of a master's thesis.

Submission date of the thesis to Institute of Graduate Studies: ___/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Dhuha Abd ALI

Signature

DEDICATION

Let me start by saying how grateful I am for the guidance and encouragement that I received from my advisor, professor Dr. Hasan Hüseyin Balık, throughout this thesis project. Thanks to him, I have learned a lot about how to improve my work. Please accept my sincere thanks to everyone who has assisted me in dispelling my fears and providing moral support throughout this master's thesis. Thanks to the unending support and inspiration of my family, I will be able to continue my journey. It would not have been possible to complete this project without their help.

ABSTRACT

OPTIMIZATION SMART CLASSROOM USING FEATURE SELECTION IN ACO ALGORITHM

Abd Ali, Dhuha

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Prof. Dr. Hasan BALIK

Date: December/2022

Pages: 73

Artificial intelligence, as we know it, is the ability of digital devices and computers to mimic and resemble the actions of intelligent individuals in several ways. One of its most important branches is that of machine learning. The internet of things, which includes smart devices and machine learning technology, provides a wealth of useful tools for humans, including smart learning or smart education. Smart education especially smart classroom is a concept that have high important with high priority specifically with the huge development of the technology and computer science, when the university or any learning foundation have the correct situation for teaching or learning the operation of teach students will be more efficiency and more generous, so this field must be receiving care by make the campus or classroom suitable for teaching. Most of the smart campus had data set that had been collected by sensors and these data sets could be taken to make filtering on them by one of machine learning technologies (feature selection) to get best results of the suitable conditions for successful learning. We used the ACO algorithm to get these results by merge it with another classifier algorithms such as (random forest, logistic regression and so on).

Keywords: ACO, IOT, FS

TABLE OF CONTENTS

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

AI	:	Artificial Intelligence
IoT	:	Internet of Things
ML	:	Machine Learning
FS	:	Feature Selection
ACO	:	Ant Colony Optimization
PSO	:	Practical Swarm Optimization

1. INTRODUCTION

We must certain that digital technologies have been an important internal part of our life, how we deal with data and information, how we look like to the information is changed, how we communicate with each other, so, because of the massive growth of these technology in the last century, most sectors of life (healthcare, agriculture, manufacture, etc.) are affected and more of challenges had found, one of those important sectors are education. The education connects and belongs to this massive development to make genuine changes in the heart of teaching and learning, so it always has new challenges [1]. Smart classroom is one of these challenges. The classroom is the basic structure of a university campus or any educational building. The classroom is a learning space that consists of tables, chairs, a blackboard, some electronic devices, lighting, and traditional things that require human intervention to manage them [2]. Therefore, the study in the traditional system depends on human effort through turning on or closing electronic devices or using traditional methods of teaching, traditional examination systems, and many things that depend on human intervention. All of these may lead to waste energy, money, time, and effort [3]. Last years the continuous developing of technology led to improve the educational operation by using smart devices or by making smart classrooms. One of the most important concepts that is widely using in smart campus is artificial intelligence (AI), the history of the artificial intelligence is more than just a record of how machines have aimed to mimic or replace a fixed idea of human intelligence. As a result, contrary to popular belief, in ways that confound teleological theories in which symbolic AI arise naturally and inexorably from attempts over ages to reduce human reasoning to a logical formalism [4]. Machine learning is one of the artificial intelligence developments which includes numerous algorithms and modeling tools that can be applied to many several types of data processing tasks. The overall purpose of machine learning (ML) is to find patterns in data that can help solve problems that aren't immediately apparent such as the handling with the huge data that is produced by sensors in smart classroom [5]. Smart technology means adopting and developing modern technologies that produce new products, shorten development periods, achieve resources efficiency, or deliver more personalized services. Smart technology is being put to a variety of uses, including in "smart" campuses, "smart" phones, and "smart agriculture.”, etc. [6].

The concept of information and communication technology (ICT) exploited by smart campus for producing the improving, growth and development smart solutions, and this will make many benefits such quality live and education improving [7].

college students on campus will be given access to information about their environment and neighboring buildings to better their work and study time. students can use sensors to monitor the conditions of various study areas across campus, allowing them to choose the optimum place to study at any given time. the committee agreed that temperature, pressure, and humidity were the most essential criteria in determining a suitable study setting [8].

The "Smart Classroom" is a cutting-edge educational tool that utilizes the latest in computer technology. It is a brand-new approach to educating students using technology in the classroom. For the time being, most of the study of computer-assisted instruction takes place in the classroom. Learning in a smart classroom has become the primary vehicle for implementing wisdom education. It is also the focus of contemporary school informatization teaching reform and product research and development. Teaching reform and new product development in education technology are centered on this area as well. When it comes to conceptualizing or building a smart classroom in practice, a universally applicable template does not exist. [9].

When it used in this way, smart classrooms can help transform traditional education into a more contemporary model that increases student achievement while also expanding student access to educational opportunities. Smart classrooms use a wide range of computer hardware, such as computers, tablets, interactive whiteboards, smartphones, and other mobile devices. Teachers are faced with classroom issues such as excessive talking during instructions, getting out of seat without permission, throwing objects across the room t, sleeping during classroom instructions and disrespect to the teacher. It is important that teachers find creative ways to deal with the issues as well as provide quality instructions in the classroom, smart classroom could handle with these problems [10].

With its significant development and wide adoption in a variety of domain applications, the Internet of Things (IoT) means the way people and objects interact with each other. It is via the integration of the functionality of contiguous IoT nodes that these linked devices, also known as IoT nodes, come together and work as a team. The rise of large-scale IoT sensing networks has been accompanied by this trend, and they have become the infrastructure for a wide range of applications [11].

Teachers and students can now communicate with each other in new ways thanks to the Internet of Things, which has impacted the way people and objects interact with each other and the teaching and learning process. Students can now employ modern technology in and out of the classroom to complete assignments and instructional activities. The Internet of Things (IoT) is on its way to becoming a pervasive global computing network that will connect everyone and everything to the Internet due to the rapid improvements in technology. Teachers and students can benefit from the use of IOT in education.

Knowing that education is powerful avenue for social evolution, social advancement, and helps create and construct younger members of our country's population, the teacher's use of technology during the learning sessions is critical to the manifestation of this power. The Internet of Things (IoT) represents cutting-edge technology, yet new tools for technological progress that make up the IoT is not something we have not encountered before [10]. Campus comfort and energy efficiency are monitored by IoT sensors, which then take appropriate action [12].

Many industries, including autonomous driving, health care, banking, manufacturing, energy harvesting, and others, have experienced a rapid surge in the use of machine learning (ML) techniques during the last decade. Much like computers in the 1980s and 1990s, machine learning is often seen as one of the most revolutionary technologies of the present [5].

The term "machine learning" encompasses a wide range of techniques, from fitting prediction models to data to finding patterns in data. A branch of computer science concerned with learning from data equivalents or mimics of human behavior recognition of patterns, albeit objectively computed reasoning This is especially true in machine learning. A time saver when dealing with a huge dataset (many discrete data points) too complicated (includes) human analysis and/or a huge number of features in cases where data analysis is to be automated to create a repeatable and time-saving process [13].

Many discrete data points or many data sets could be treatment by feature selection because to be successful in the field of machine learning, it is essential to master the art of feature selection. A primary goal of the feature selection problem is to keep performance accuracy while shrinking the feature collection [14].

Most of the time, machine learning is employed to solve problems with a plethora of features. As a result, finding an ideal collection of features and eliminating unnecessary ones is difficult. Some features in a dataset are irrelevant because they are redundant or irrelevant in some other dataset. As a result, considering such characteristics is counterproductive and usually results in subpar classification accuracy. As a result, feature selection aims to improve classification efficiency by picking only a small subset of acceptable characteristics from the initial vast range of features. Remove redundant and irrelevant features to reduce data dimensionality and speed up the learning process [15]. The reason behind using this topic is because my work in education sector, so the project major motivation is to produce a suitable smart campus or classrooms for students in an appropriate environment, this environment comes true by providing the appropriate conditions. the conditions are consisting of measuring (temperature, humidity, lighting, pressure, noisy and movement). the big data that will be collected from the sensors in smart campus will be filtered and processed to get the best results. these results will help to make correct decision for appropriate environment. the project has as its goal the implementation of an entire system on a college campus. As an added benefit, the creation of a Global Connector has been considered, which would allow for the flexible and effective integration of various platforms. the sensors that generate the data; four different platforms for integrating and managing the data; and a third platform for storing data. Two components: a Server, which stores and publishes data, and a Connector, which is a software program that connects the server to other systems.

1.1 PROBLEM STATEMENT

The major problems in WSN based IoT in smart building is the huge amount of the data set that collected by wireless network sensors (WSN) with the using of MQTT protocol and how this data could be filtered to get the best results, This thesis addresses the importance of reconsidering the data that comes from sensors in smart classrooms by selecting only the features that are expected to be more important than the other features in our calculations. Selecting key features from a larger set of features will help reduce the number of dimensions and hence fighting the curse of dimensionality issue. As such, in this study, what we suggest using one of the popular optimization algorithms, called Ant Colony Optimization (ACO), along with another popular classifier, which is Logistic Regression (LR), to build a hybrid model. The proposed hybrid model, which was built based on a series of experiments on two sensor datasets, will help us select the appropriate set of features in smart classroom environments.

1.2 RESEARCH QUSTIONS AND OBJECTIVE

a. What is the smart classroom? What the IOT components?

b. How can we use Ant Colony Optimization (ACO) algorithm and machine learning for feature selection in smart school environments?

c. How can we check the efficiency of the proposed ACO-LR model?

d. How can we build a classification model based on ACO-LR model to use on other unseen data?

1.3 OBJACTIVES

a. Explain the concept of smart classroom, IOT, and the importance of them.

b. We will build a hybrid model of ACO algorithm and Logistic Regression (LR) classifier, called (ACO-LR), which can select best features in smart school environments. Choosing LR for this task is based on several experiments.

c. We will compare the performance of ACO-LR model against other optimization methods like Practical Swarm Optimization (PSO) and Genetic Algorithms (GA). The process will be done on two datasets of different size.

d. We will use the best selected features and based on experimenting with several popular classifiers, we will use the best classifier to build a classification model that is ready to run on unseen data in smart school environments.

1.4 TARGET GROUPES

The target groups of this thesis will be three groups. Firstly, the researchers because they will find good model for feature selection and classifier, secondly the student because they will find appropriate situation for education and learning, and thirdly teachers, and the field of education. Because of the utmost importance of education operation in our live the buildings (schools, universities, and classrooms) must become comfortable, and all appropriate conditions must be available.

1.5 CONTRIBUTIONS

This thesis explains how the smart classroom using an IoT system-based MQTT method to measure different electrical devices in the classroom, such as air-conditioners, smart boards, intelligent test systems on smartphones, fans, light, humidity, temperature, etc. Naturally, these devices had a massive amount of data, so we propose the feature selection operation to get the clean and useful data to take the advantage of this data by merge ant colony algorithm (ACO) with several classifier algorithms to make the best feature model. The model that we will produce it will be evaluated on other data sets to make sure that our procedure is working correctly. Then we will take the best selected feature with some of classifier's algorithms to build classification model that is ready to use.

1.6 THESIS OUTLINE

Thesis structure is arranged as follows:

A. Chapter 2 the element of the work is explained in detail.

B. Chapter 3 contains the proposed model, results, discussions, and suggestions from the actualization of the proposed model.

C. Chapter 4 is for the results of the work and conclusion.

D. Chapter 5 demonstrates our outcomes and presents suggestions

2. BACKGROUND

2.1 INTRODUCTION

An overview is provided in this chapter about the elements that will be used in our work. This work is about the smart campus or smart classrooms. We will write about the concept of smart classroom and what is the importance of it in the education sector (the advantages). We will explain the term of (IOT), WSN, and MQTT because they are used in the structure of the smart campus building. The algorithms that are used in our work to oversee with our data set to get the best results by optimization and calcification data will be explained, and the term of machine learning, the types of it, the most important algorithms in machine learning will be explained, and what are the applications of it.

2.2 SMART CLASSROOMS

Smart classroom is a closed environment technology that it used for improving teaching and learning [16]. The exemplary smart classrooms should have tools or devices for variety logical duties such as smart board or smart cameras. Using smart classrooms means more intelligent students, more professional education environment, so students can be great in informational social environment and encourage the development [17].

2.2.1 Smart Classroom Advantages

There are many advantages of using the smart classrooms especially when compared with traditional classroom, there are some of these advantages [18].

A. Learning ideas: - Classroom activities are planned according to instructional design to assist students in learning latest information and reviewing previously learned information

B. Learning methods: - Students become the focal point of knowledge acquisition, with the classroom serving as a laboratory for evaluating the impact of knowledge acquisition and practice

C. Learning sources: - students become energetic constructors in the operation of learning, combining their learning habits and learning abilities

D. Evaluation of Learning: - through three steps the smart classroom primarily regulates the students in each of through three situation (per-class, in – class, after – class) of incorrect question analysis and class interconnection as well as the application of data diagnosis, an intelligent learning system can easily be seen.

2.2.2 The Smart Classrooms Components

The architecture of smart classrooms consists of several of elements such as Internet-in-the-Cloud, Mobile Device, and Wireless Connection real-world situations. All these elements will produce a huge amount of data [19]. The key to smart campus is represented in four aims [20]

A. devices and sensors availability.

B. ability of gathering, storing, and make operations on data.

C. getting how and when data should be used.

D. set of benefits for using data.

The encouraging standard is famous as the "Internet of Things" (IoT). Its growth works hand by hand with forward movement of helping mechanization labeled to this novel concept of the wireless transmission script, like as wearable sensors, RDF, WSN, actuators, Machine-to-Machine(M2M) methods etc. [21]. The jobs of the IOT could be resolved within the next essential components: Association, Sensing, counting, and Assistance the wireless transmission scenario.

2.3 IOT (INTERNET OF THINGS)

The Internet of Things (IoT), which symbolizes a group of connected devices, has emerged because of the rapid expansion of information and communication technologies (ICT). Wireless links (Wi-Fi, RF, NFC, Zigbee, etc.) are used to exchange data between things and people [22].

2.3.1 Identification

The Internet of Things (IoT) offers capacity and elasticity to quickly acclimate to changing conditions. The programs deployed on them make the environment intelligent. IoT is made up of both hardware and software [23]. The basic goal of IoT is to make devices able to be always connected and from wherever via a lattice. In IoT, thousands of smart objects, sensor's nodes, and applications in real-time build up and engage data to determine time. IOT is using in different fields like industry, agriculture, healthcare, and education by making smart classrooms [24].

The interconnection and repetition among objects need requirements counting electrical mechanisms, developing technologies and procedures for uniquely recognizing the several devices over the network. To this end, there is comprehensive collection of possible identification techniques, and they could be categorized into three principal classes [25]:

A. Thing Identifiers are utilized to recognize environmental and practical devices uniquely. Electronic Product Codes (EPC), Universally Unique Identifiers (UUID and so on from this group.

B. Connection Identifiers are utilized to recognize methods pending the interaction with other things, counting internet-promoter transmissions. Classical models of connection protocols are the IPv4, IPv6 addresses.

C. Application protocols are techniques to serve Application-Status IDs. In the range of IoT, the description of reinforcements and assistance includes Uniform Resource Identifier (URI), Uniform Resource Locator (URL).

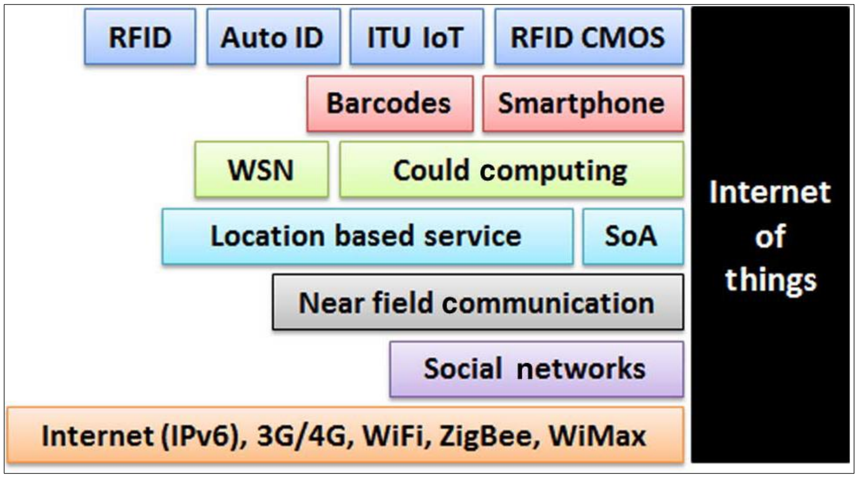

Figure 2.1: TECHNOLOGIES ASSOCIATED WITH IOT [23]

2.3.2 Sensing

The sensing of IOT component distributes with two prime missions:

A. collecting and assembling conditions data from things within the network.

B. transferring data as mentioned above to the database or a knowledge warehouse.

The processing and interpretation of this knowledge will enable producing judgments and designations depending on the service. For example, various mobile credentials would allow users to control and manage a significant number of intelligent machines operating each aspect of their house (i.e., convenience, protection, and power exhaustion) [26].

2.3.3 Connection

For IoT purpose, sensors or more common, the connections, should communicate and interact across a mechanism (i.e., the wireless incidence channel) influenced by noise, extinction, and the protection problems. The purpose of transmission methods is to facilitate the association and interaction of various things to implement a secure, quick, and guarded brilliant service. The list of connection techniques selected in the IoT scenario has Wi-Fi, RFID, NFC, Bluetooth, moderate power variant, and ZigBee. To a complete explanation of the characteristics of these intelligence techniques (i.e., engaging band, productivity, variety of communication, and data encryption/protection tools) [27].

2.3.4 Counting

The counting ability of IoT is realized by the handing systems like GPU, SOCs, and software purposes. On the business, various implements stand was produced to operate with IoT employment. Some tools principles IoT- specified are Beneficial Arduino, ARM and Raspberry Pi. Furthermore, several series were revealed to accommodate IoT functionality. In these circumstances, OSs are necessary because they operate for the complete lifecycle of the machine. Therefore, in specific, the Real-Time Operating Systems (RTOS) are essential competitors in the field of IoT. Furthermore, the Cloud platforms play another critical role in the computational assignment of IoT. These present an assortment of settings. [28].

2.3.5 Services

IoT services could be represented by four classes. (Identity-concerning Services, data gathering service, Cooperative -Aware assistance, and broad spreading Services) which are presented in (Gigli and Koo (2011) and Xiaojiang et al. (2010)). Sets involved in original class (identity-concerning Services) are used for all IoT purposes. They proposed at recognizing real-world things to produce them in practical environment. Consequently, an Identity-concerning set consists of the items provided with a mechanism of association (i.e., RFIDs, NFC) and a mechanism that holds status of presented thing.

Data gathering service interest with the processing of collecting data from various sensors positioned in the context. Parsing and normalization are standard procedures worked on producing new data. Furthermore, specific settings are dependable for communicating and announcing data through system base to the IoT utilization.

The cooperative -aware assistance supports the aggregated data collected from the prior settings to obtain conclusions and complete responses. The effectiveness of aid relies on the safety/speed of the IoT field and computation ability of the tools. [29].

2.3.6 Internet of Things Construction

Any building blocks of an IOT system consist of several layers to represent its functions. the bottom layer consists of different sensors, and they are linked to each other to form a sensor network. the sensors are linked to a central unit called the gateway, which is an intermediate device through which the signals produced from the various sensors are collected and processed. the gateway is connected to a higher layer by various wireless networks such as GSM, 4g, or LTE networks. the final layer consists of cloud computing to store and analyze signals [30]

2.4 WIRELESS SENSOR NETWORKS

Wireless sensor networks (WSNs) are applicable to a variety of industries, including agriculture, health care, and the energy sector. Multiple heterogeneous sensor nodes can collaborate to accomplish a particular task in a particular area of interest [31].In a typical WSN, sensors are used to monitor and record the physical and environmental conditions (temperature, humidity, radiation, lighting, sound pressure, etc.) of an area (smart classroom). Together, these sensors collect data and transmit it to a central location via the network. This technology can facilitate the delivery of distance learning courses (Memos et al. 2020). Homogeneous and heterogeneous WSNs are categorized based on the various sensor types employed. The homogeneous sensor network and the heterogeneous sensor network are both referred to as sensor networks with a wide variety of sensors. When it comes to sensor routing protocols, WSN prefers communications that are both dependable and efficient. Chargers are provided for the nodes. [32].

There are several sensor nodes that BS is interested in monitoring. Analysis and reduction of data similarities are used for decision-making. Even more importantly, these data can be used locally, but they can also be sent to other networks located in remote areas. Sensor nodes, on the other hand, would be unable to deal with such a high communication overhead.

Data aggregation is a term used in WSN for the purpose of describing the processes of aggregating data from many sensors and reporting it to BS. Clusters of sensors are used. The cluster head is the most important node in each cluster (CH). To be transmitted to a central station, each CH gathers data from its sensor nodes (BS). Sensors have been deployed with non-replaceable battery-powered sensors. The amount of energy used is crucial concern for the welfare of the World Wide Web. In WSN, batteries are critical because they serve as an indicator of how long the network will last. Data transmission consumes most of the energy in a wireless network. As a result, protocols for energy-efficient routing are required. [33].

Nack explains that a wireless sensor network form is a predetermined number of autonomously working devices, each of which is outfitted with sensors for inspector physical or environmental features. There is a logical structure to the devices, which are all connected. There are a variety of ways to structure a wireless sensor network (WSN). single-hop star: immediate connection between sensor sections and gateway; numerous -hop mesh: a mesh with sensor sections and a gateway; two-tier pyramidal cluster: many clusters, each with several sensor nodes and a gateway; It is up to the WSN's purpose and context to determine how many devices are connected to it.

Djedouboum et al. conduct a lengthy discussion on WSN architectures, applications, and challenges as sources of big data [34].

2.5 THE MESSAGE QUEUE TELEMETRY TRANSPORT PROTOCOL

It has emerged as a promising technology, the Internet of Things (IoT). The IoT devices' limited resources have resulted in data transfer limitations. To meet these demands and limitations, new protocols have been proposed. IoT application protocols such as MQTT, Constraint Application Protocol (CoAP), and many others have been proposed in the past few years [35].

Communications between devices in an IoT environment are oversaw by MQTT, a protocol Developed by IBM, the protocol is a machine-to-machine communication method. A third type of reliability mechanism, referred to as Quality of Service (QoS), can be selected by the protocol to ensure secure transfers (QoS).

Environment is ideal for this model Transport layer protocol TCP/IP is used by MQTT (Message Queue Telemetry Transport), an application layer protocol. Andy Stanford-Clark and Arlen Nipper were the original developers of the protocol. In IoT systems with limited resources. Communication takes place in a publish/subscribe fashion in which the broker oversees delivering messages to their subscribers. Clients for MQTT. As a result, neither the publisher nor the subscriber must be aware of the other's existence.

The rules and regulations of conduct This standard has been defined in ISO/IEC 20922 and is presently an OASIS (Organization for the Advancement of Structured Information Standards) standard [36].

2.5.1 Message Queue Telemetry Transport Advantage

Following the advantage of (MQTT) protocol [37].

A. The MQTT design has a significant advantage in terms of scalability. It allows many small, constrained devices to communicate asynchronously, making it an ideal solution for mesh networks.

B. Messages can have three types of various levels of quality of service (QoS) support: "fire-and-forget/unreliable," "at least once," and "exactly once."

C. the protocol is widely used because of its characteristics such as lightweight, low bandwidth requirement, openness, and ease of implementation, all of which make it ideal for IoT systems trust worthiness

2.5.2 Message Queue Telemetry Transport Components

According to [38] Message Queue Telemetry Transport consists of three parts:

A. A programmer or a creator (An MQTT Client)

B. a trader (An MQTT Server)

C. Subscriber/user (An MQTT used)

To use MQTT, a client must be installed on a computer or other device. Open network connecting to the server, publish messages, subscribe to request messages that interest you, unsubscribe from requests that you no longer want to receive, and close network connection to the server are all responsibilities of the client. The MQTT protocol's application message is the data that the application sends across the network via the MQTT protocol. Message load data, a Quality of Service (QoS), a combination of qualities, and a matter name are all included in MQTT application messages. An MQTT server is a program or system based on MQTT that serves as an intermediary between subscribers and publishers.

Accepting network connections from clients, admission application client-posted messages, handling subscriber and unsubscribe orders from clients, delivering application messages to clients in accordance with its participation, and shutdown the network connecting from the client are all tasks performed by a MQTT broker. There are two ways to communicate with MQTT. This facilitates data exchange as well as device management and control. There are no restrictions on the size of the message payload if the XED header is two bytes or less.

2.6 MACHINE LEARNING

People have pushed themselves to the limit of thinking in this digital age by integrating artificial and human brains. By putting up a show. As a result, a new industry has emerged. It is referred to as AI. Artificial intelligence (AI) is an approach that is based on prior knowledge and so aids in prediction by revealing the precision with which future values will be. So, data is processed using an algorithm. There are a lot of them. Machines, deep learning all fall under the umbrella term. As a result of machine learning's growing prominence, values have risen recently. Resurrection has only occurred recently. Implementers are now competent to deliver in this area better disclosure of the algorithms' workings. ML, one of the fields of artificial intelligence, numerous new possibilities have opened. ML is built on the foundation of arithmetic and statistics. Object distinction is one of many uses for machine learning. Speech recognition and text classification are also included, forecasting the weather, and examining ruined objects Food production, facial recognition, medical diagnostics, and other applications are examples of this. All the above are result of ML's reliance on data, Big Data is necessary [39].

played an important influence in its development in our daily lives everything is now digitally registered. Data from the Internet of Things, cybersecurity, smart cities, the business sector, smartphones, social media, health, and a slew of other sources abound in today's technological environment. In species of Real-World Data and Machine Learning Techniques, data could be categorized as structured, semi-structured, or unstructured; these types of data are on the rise.

Data mining can be used to generate insights for a wide range of intelligent applications in the relevant fields. relevant mobile data can be leveraged to construct personality context-aware smart mobile implementation, as well as to build an automated and intelligent cybersecurity system. As a result, real-world applications rely on data management tools and processes that can quickly and intelligently extract insights or meaningful knowledge from data [40].

When Samuel2 produced the idea of machine learning in 1959, he did not expect it to be used in a variety of fields, such as computer vision and general game play. With the advent of artificial intelligence and machine learning, researchers in the mainstream AI sector as well as professionals in other fields are making significant advancements in their respective fields.

C60's solubility was first detected using machine learning 12 years ago, and it has since been used to identify new materials, forecast material and molecular properties, to research quantum chemistry and to build pharmaceuticals. Machine learning resources and technologies become more widely available, the barrier to entry for using machine learning in materials science has never been lower [41].

Machin learning (ML) seeks to automatically identify significant relationships and patterns from examples and observations rather than codifying knowledge into computers. ML advances have capable the recent rise of intelligent systems with human-like cognitive capacity that penetrate our business and personal lives and shape the networked interactions on electronic markets in every conceivable way, with companies augmenting make decision for output, correlation and employment retention, trainable support systems adapting to user predilection and tradable support systems adapting to user preferences [42].

2.6.1 Applications of Machine Learning:

There are several applications of machine learning in our live such: - [43][40].

A. Intelligent decision-making based on predictive analytics: Using data-driven predictive analytics to make intelligent decisions is a prominent use of machine learning. Using historical data and correlations between explanatory variables and anticipated variables, predictive analytics attempts to make educated guesses about the future.

B. In today's digital age, it is critical to safeguard your networks, devices, and data with the help of threat intelligence, which is a branch of cybersecurity. An important part of cybersecurity is the using of artificial intelligence (AI) and machine learning (ML), both of which constantly learn by data analyzing to spot manners, better detection malware in encrypted traffic, and uncover insider risks.

C. Smart cities and the Internet of Things (IoT) a crucial part of Industry 4.0, the Internet of Things (IoT) is a technique that allows common things to transmit data and conduct tasks. Consequently, the Internet of Things (IoT) is seen as the major frontier that has the potential to improve every aspect of our lives. An important application area for the Internet of Things (IoT) is smart cities, which use technology to enhance city services and the quality of life for people. Machine learning is a vital technique for IoT applications because the using of expertise to spot trends and construct models that may prediction behavior in the future and occurrences.

D. All nations must have a well-developed transportation infrastructure to grow economically. Many cities around the world, however, are witnessing a significant increase in traffic volume, which has resulted in critical difficulties such a delay, traffic congestion, higher fuel prices and increasing CO_2 pollution. Consequently, a smart city requires an intelligent transport system that can predict future traffic, which is a necessary component. Minimizing problems with traffic flow can be made easier with machine and deep learning modeling-based traffic prediction.

E. Several medical fields, including disease prediction, medical knowledge extraction, finding regularities in information, patient administration, and so on, can benefit from machine learning. According to the WHO, a recently discovered coronavirus is responsible for Coronavirus Disease. In the fight against COVID-19, learning strategies have recently gained popularity.

F. In the world of e-commerce, product suggestion is one of the most utilized uses of machine learning, and it is a significant feature on every e-commerce website today. Customers' purchasing history can be analyzed with the support of machine learning technology, and firms could make personalized product recommendations based on their behavior and interests.

G. Natural language processing (NLP) and sentiment analysis: NLP involves the reading and interpretation of spoken or written language through the computer. As a result, computers, for example, can read a piece of text, listen to a piece of speech, and decipher its meaning, as well as assess its sentiment and determine which components of it are most important, using machine learning techniques. Some examples of NLP-related tasks include virtual personal assistants, chatbots, speech recognition, document descriptions, and linguistic or machine translation.

H. Images may be recognized as digital images, speech can be recognized as text, and patterns can be recognized as images A well-known and often used machine learning technique is image recognition. Image recognition could be used for a variety of purposes, including determining whether an x-ray is malignant or not, identifying characters or faces in an image, and suggesting tags on social media, such as Facebook.

Voice-activated assistants such as Google Assistant (GA), Cortana (CA), Siri (SIRI), and Amazon's Echo (ALEXA) use machine learning.

I. Agriculture is crucial to the existence of all human endeavors, and it must be sustainable. Improved yields and reduced environmental effect are two benefits of sustainable agricultural practice. Supply chains for sustainable agriculture rely on information, skills, and technologies such as the Internet of Things (IoT) and mobile devices, as well as knowledge transfer, which encourages farmers to embrace sustainable agriculture techniques. For example, machine learning can be used in the pre-production phase to predict crop yields, soil properties and irrigation requirements; in the product phase for weather prediction, diseases detection and weed management; in the processing phase for demand estimation, production planning and distribution; and in the distribution phase for inventory management, consumer analytic.

J. Analytics of user behavior and context-sensitive mobile applications: Knowledge of its surrounds at any given time is what is meant by the term "context-awareness". Computers with context-aware capabilities gather and analyze data on their own, without the need for human intervention. Machine learning approaches, which can learn from contextual data, have had a significant impact on mobile app development thanks to the potential of AI. It is thus possible to employ machine learning to design mobile applications that can understand human.

2.6.2 Machine Learning Algorithms

There are many algorithms that are used in machine learning, here some of these algorithms: -

A. Support vector machine: - support classification and regression problems can be overseen with SVM. The decision boundary for this method is defined by a hyperplane. Decided plane is needed when there is a group of items that fall into distinct classes. Class members may or may not be linearly separable, necessitating the use of sophisticated mathematical functions called kernels to separate them. Based on the examples in the training data set, an SVM seeks to accurately identify the items [44], [45].

B. XG Boost Algorithm: - It is called Extreme Gradient Boosting, or XG-Boost. Regression and classification predictive modeling are two of the most common uses for GBM, which is a type of Gradient Boosting Machine (GBM). GBM models regularly outperform other machine learning algorithms, as seen by their superior results on a variety of machine learning benchmark data sets. As part of an ensemble method, XG-Boost creates new models to correct the residuals or errors of previous models before combining them with one other to forecast the result. XG-Boost has been shown to be faster than s XG-Boost is the most extensible and adaptable of the Boosting Tree models. The learner model is strengthened by the incorporation of numerous tree models. Parallel computation can be sped up by XG-Boost's ability to take advantage of the CPU's multithreading capabilities [46].

C. Gradient Descent: in this algorithm the goal is to find the least expensive path. In theory, it should be able to compute a partial derivative of the function that is the slope or gradient. By subtracting the negative derivative from the coefficients at each iteration and reducing each step's learning rate (step size) multiplied by derivative, it is possible to obtain local minima after only a few iterations. When the cost function converges to the least value, the iterations are stopped and there is no further reduction in the cost function possible. SGD (Stochastic Gradient Descent), BGD (Batch Gradient Descent), and MBG (Mini Batch Gradient Descent) are all variations of this technique (MBGD) [45].

D. Multivariate regression analysis: - There is only one independent variable in simple linear regression models. Problems in the actual world, on the other hand, are much more complicated. In most cases, more than one factor affects a single dependent variable. It is a many-to-one relationship between independent (predictor) and dependent (output) variables in multiple linear regression. There is no guarantee that a regression will perform better or make better predictions just because additional input variables are included in it. No matter which type of regression you employ, it's not better than the other [45] [44].

E. Bayesian learning: it is possible to use the previous posterior distribution as a prior when new observations are available. The Bayesian network can oversee incomplete datasets. The method prevents data from being over-fitted. Attempting to remove inconsistencies from data is not necessary [44].

F. logistic regression algorithm: To examine the relationship between a binary or outcome and multiple contributing factors, logistic regression is a type of multiple regression method that includes multiple of logistic regression, conditional of logistic regression, polytomous of logistic regression, ordinal of logistic regression, and adjacent categorical logistic regression [47].

To classify data, logistic regression is the most employed method. Because it provides the binomial probability, it provides the binomial outcome Whether or an event takes place is determined by values (in terms of 0 and 1). a set of parameters to be entered [44]. In terms of binary classification and predictions, logistic regression has been widely employed as a broad data processing tool. It does not have data points grouped in a linear fashion. These data points are all grouped together in a single pile with a single category label as shown

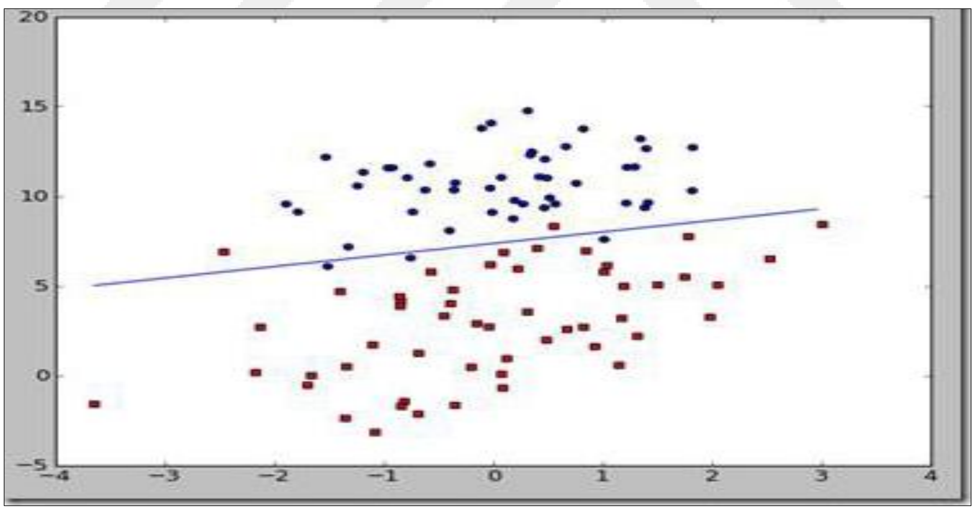

Figure 2.2: Logistic Regression [48].

Following this initial training phase, a predictive model is built using the previously collected values of predictor variables (called covariates) and related outcomes of interest. Next comes testing to determine whether the model's accuracy was improved after it was trained. To accomplish this, the dataset is separated into training and validation data sets. The model's output is compared to the known outcome and evaluated considering the provided covariates. Models can be used to predict future outcomes if their classification is consistent with test results for most data sets, and this can be done by reevaluating their model for new measurements of the variables.

As a result of its simplicity and versatility, logistic regression is often used to address a wide range of issues. In medicine, logistic regression is used to estimate a patient's likelihood of contracting a specific disease. Personal information like as age, income, gender, place of residence, and previous votes can be utilized in politics to make educated guesses about a person's voting intentions. Logistic regression can be used in finance to forecast the risk that a homeowner would default on a mortgage or that a credit card transaction will be fraudulent. Logistic regression, like all machine learning technologies, necessitates a substantial amount of training data to be effective.

The logistic regression model can be summarized in the following 2.1 equation: -

$$\mathbf{Pr}(\boldsymbol{y} = \pm\mathbf{1}|\boldsymbol{x}, \boldsymbol{B}) = \boldsymbol{\sigma}(\boldsymbol{Y}\boldsymbol{B}^{\boldsymbol{T}}\boldsymbol{x}) = \frac{\mathbf{1}}{\mathbf{1}+\boldsymbol{e}^{(-\boldsymbol{y}\boldsymbol{B}^{\boldsymbol{T}}\boldsymbol{x})}} \qquad \textbf{(2.1)}$$

Since the model parameters are presented by the vector $\beta=(\beta_0,\dots,\beta_d)$, y the class label and the vector x=(1,x1,…,xd)∈Rd+1x=(1,x1,…,xd)∈Rd+1 would be the covariates [49].

There are many Advantages of Logistic Regression such as [45]

A. a low barrier to entry

B. efficiency in computation

C. from a training standpoint

D. regularization is a snap

E. Binary Logistic Regression: -

To model the link between many continuous or categorical independent factors and a dichotomous dependent variable, binary logistic regression is utilized. Binary logistic regression has various assumptions that must be met to produce an accurate result: -

A. The answer variable's logic should have a linear relationship with the explanatory variables.

B. There should be no correlation between errors.

C. To avoid multicollinearity, the explanatory variables should have low correlations.

D. There should be no outliers, high-leverage values, or spots with a lot of sway.

When using logistic regression, it is presumed that the explanatory variables will have low correlation. The assumptions of the logistic regression model must be met for the findings to be considered legitimate. Invalid statistical conclusions may be drawn because of model flaws such as excessively inflated standard errors, erroneously low or high t-statistics, and parameter estimates with nonsensical signs. This is not true for observational data, where the explanatory variables cannot be orthogonal to each other in experimental design. Non-experimental scientists have long recognized that data collinearity is an inherent and painful aspect of life due to the unpredictable activities of the data generation process. Variables that are highly linked are frequently gathered and analyzed in surveys [50].

2.7 KINDS OF MACHINE LEARNING

Machine learning has four types (supervised, unsupervised, semi-supervised and reinforcement) learning. When the data is in the form of input variables and output target values, supervised learning is used. Learning the mapping function from input to output is the primary goal of the algorithm. Large-scale labeled data samples are readily available, making this an expensive approach for applications with a limited amount of data [51].

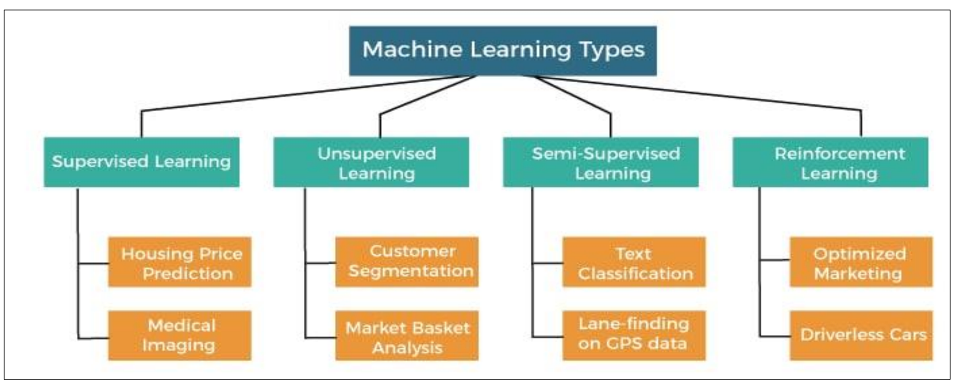

Figure 2. 2: ML Types

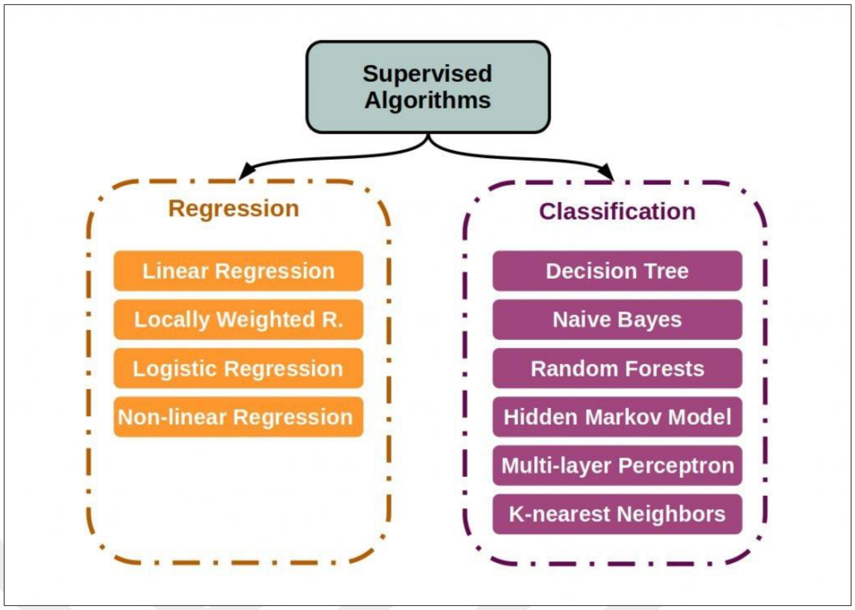

Figure 2.3: Supervised Algorithm

A. Supervised learning There are two primary types of strategies (classification, and regression) to consider [51]

I. Classification

Classification ways have been used in a wide range of scientific domains. To develop a classification model, training data are utilized to predict the class label for a new sample using an algorithm. It is possible to have a discrete or continuous output from a classification model, such as the decision tree classifier. However, the results of learning algorithms must be thoroughly examined and analyzed, and this interpretation must be accurately understood to compare various learning algorithms.

Scalar values like precision, responsiveness, and resolving power are used to measure classification performance. Comparing classifiers with these metrics is simple, but it comes with a slew of drawbacks, including a high sensitivity to data imbalances and the omission of some classes' performance. Different interpretations of categorization performance can be derived from graphs such as receiver operating characteristics (ROC) and precision-recall curves [52].

ML algorithms can describe any instance in any dataset by describing it with the same set of attributes. It is possible that the attributes are categorical, continuous, or binary. When cases are recognized and correctly labeled, then the learning process is referred to as supervised. In machine learning, supervised- learning is the process of discovering a function from training data that has been categorized. Using the results of the testing, a derived task is generated that may be used to map new examples. However, each data input object has a pre-assigned class label. Supplied data is used to train supervised algorithms, and the goal is to find a model that performs well on both known and unknown datasets. Classification algorithms are designed to accomplish this goal [53].

Prior to running the machine learning algorithms on the data, pre-processing is critical to achieving appropriate categorization results. Discrete data, elimination outliers and noise from the data, integrating data from diverse sources, dealing with missing data, and transforming data to comparable dynamic ranges are all part of this process. Among these, data normalization is characteristics must be transformed into a common range to ensure that larger numeric values do not dominate smaller numeric values. The primary goal is to reduce the influence of features that have a bigger numerical contribution in classifying patterns. To forecast the output class of a previously unknown instance, all the patterns in the data were considered equally important. All data features contribute equally to the learning process, making it helpful for statistical learning approaches.

Machine learning methods such as k-Nearest Neighbors, and Support Vector Machines (SVM) have studied the necessity of data normalization for developing reliable predictive models, multimodal biometrics systems, vehicle classification, faulty motor detection, predicting the stock market, leaf classification, credit approval data classification, genomics, and some other application areas have all been proven to benefit from normalization [54].

The smart campus used the classification so when it comes to educational data, there are several information systems in use at the university, but they are not being put to effective use. The lack of effective methodologies and unclean data, isolation data, or low content information in those methods make it hard to use multiple sources of student performance information to make accurate predictions progress [55].

There are some problems or issues in the collecting smart campus data that need to be solving, for example [56].

a. The same data exists under many names across various systems Due to the lack of a uniform data standard.

b. It is possible that the same data may be inconsistent in multiple systems because it has not been updated or synchronized in a timely manner.

c. A vast amount of data is generated by information systems every day, but only a small percentage of this data is relevant. Extracting useful information from this data is extremely difficult now.

Resolving these issues can help teachers get meaningful feedback, assist teachers forecast student performance, and provide useful support for making decisions. It is frequent practice for current techniques of student accomplishment analysis to rely mostly on the past performance of students. Student consumption and Internet access are also major predictors of student accomplishment, but these are not the only factors. To get the most out of educational data and make predictions about student performance [55].

Here are some of classification algorithms: -

a. Decision Tree Algorithm: To arrive at a Boolean conclusion at the leaf node, a data-separating sequence is used to define any path start at the root of the tree. Nodes and links are shown in a hierarchical representation of knowledge linkages. Using relations to organize data, nodes represent various kinds of information. Machine learning, image processing, and pattern recognition all rely on decision trees as strong algorithms. In Data Mining, DT is a common categorization model. Each tree has its own set of nodes and branches. To classify a category, a node represents one or more features, and every subset specifies a value that could be assigned to the node. Decision trees have found a wide range of applications because of their ease of use and accuracy in a variety of data formats. Classification and Regression Tree (CART), Generalized, Unbiased, Interaction Detection (GUIDE) , Quick, Unbiased, Interaction Detection and Estimation (QUEST) [53] [44].

b. Naive Bayes: conditional probability is used in this approach, which is basic and easy to understand. The model is a probabilistic table that is used in this method., which is continually updated with new training data. Probability tables are used to seek up class probabilities to predict fresh observations based on their feature values. "Naive" is a term used to describe an assumption of conditional independence. Assuming that all input characteristics are autonomous of one another in a real-world situation is impossible [45].

c. Random Forest Algorithm: Recent years have seen a surge in popularity for machine learning. There are a few algorithms more sought after than the decision tree for classifying or forecasting future events using previously trained data. An extended version of decision trees, Random Forest can predict future instances with multiple classifiers rather than one classifier to achieve greater accuracy and correctness. The Random Forest model's performance has been examined and compared to that of other classification models that produce institutionalization, regularization, connection, and a high proclivity for change, with a focus on the learning models' choice as a key feature. Many classification and regression trees are built using random training datasets and random subsets of predictor variables in the random forest setting. A prediction is made for each observation based on the results of each tree. As a reminder, RF can be used for two different purposes in practice. Developing an accurate classification or regression rule that can be applied to future data is a primary goal in some radio frequency (RF) applications. [57], [58], [44].

d. Hidden Markov model: Hidden Markov Model (HMM) and its extended variations are strong data mining algorithms that can be used to detect potential Intruders and deviant behavior. This is a possibility. Actual behavior is compared to the expected behavior typical system behavior as modeled by this. In the shadow of the Markov chain Computer systems under Multi-Stage Network Attacks have been extensively modelled (MSA), Learning, decoding, and evaluation are three of the most common issues with HMMs. An observation sequence is used to determine the HMM parameters, such as the state transition and observation probability matrices. The settings are modified so that the observation probability is maximized. As a second step in decoding, a particular HMM's state transitions are examined to ascertain the sequence of states for an observation sequence. The probability that an observation sequence was generated by an HMM is obtained at the end of the evaluation process. [59], [60]

e. K-nearest Neighbors: The k-nearest neighbor method (k-NN) developed by Cover and Hart (1967) is a basic yet effective data mining approach. In classification difficulties, this is based on the idea of similarity between samples. With large datasets, the k-NN algorithm has several challenges, including high computational costs, large store requirements, sensibility to noise and inability to work with incomplete information [61].

II. Regression

A regression model is a type of model that produces real-world values based on the input data. In a regression model, a single or many independent variables that might be quantitative or categorical and dependent variables that are quantitative are represented by a single or multiple regression lines. Classic statistical learning models, which are based on classic statistical learning theories, and deep learning models, which are primarily based on neural network theories, are the two basic types of regression models [62].

For two theories, regression is a technique. To begin, regression analyses are frequently employed for forecasting and prediction, two areas in which machine learning and regression analysis have a lot in common. Second, regression analysis can be utilized in some circumstances to identify causal relationships between the independent and dependent variables. It is important to note that regressions alone do not indicate the relationship between a dependent variable and a fixed dataset [63].

B. Reinforcement Learning

When the state and action spaces are small, reinforcement learning can be used to help network entities find the best policy, such as decisions or actions, based on their current states. When it comes to artificial intelligence (AI), reinforcement learning is the most important one study directions in machine learning. It is possible for a machine to learn by making decisions and then adjusting its strategy based on what it learns from those decisions and the effects of those decisions. However, this process of learning the fact that it takes so long to learn about an entire system makes it unsuitable for large-scale networks, even though it has been shown to converge. Reinforcement learning, as a result, has a limited practical application. It has only been recently that deep learning has been touted as a potentially game-changing innovation [51].

C. Unsupervised Learning

A computer method is used in unsupervised learning to identify patterns and subgroups in unlabeled input data [43]. Clustering and dimension reduction are two types of unsupervised learning [64]. It is common to utilize the K Means Clustering Algorithm to solve the clustering problem. An unsupervised sort of learning is involved. It may be necessary to choose between "supervised learning," "unsupervised learning," "semi-supervised learning," and "reinforcement learning" depending on the types and categories of training data available [45].

D. Semi-Supervised Learning

An intermediate between supervised and unsupervised learning methods, as the name suggests. Unlabeled and labeled data are used in the training of these algorithms. There is little labeled data, but a lot of unlabeled data. Prior to labeling new data, previously labeled data is utilized to cluster comparable samples using an unsupervised machine learning method [43].

2.8 OPTIMIZATION ALGORITHMS

Modernistic using in manufacture, business, smart education, and information systems require a tremendous amount of optimization. [65].

Parallel processing can benefit from intelligent optimization algorithms since they have global optimization performance and significant adaptability, making them ideal for this type of problem. Within a specific amount of time, these algorithms can locate the best answer or an approximation of the best solution.

There are several intelligent optimization algorithms abound. For example, genetic algorithm (GA) that simulates Darwin's notion of biological evolution by mimicking Darwin's natural selection and genetic mechanism, ant colony optimization (ACO), the differential evolution (DE) algorithm Particle swarm optimization, the shuffled frog-leaping algorithm (SFLA), bee colony, the firefly algorithm.

Machine learning researchers are also using intelligent optimization techniques. NP-hard problems such as scheduling, image, feature selection, detection, path planning, cyber-physical social system, texture discrimination, factor evaluation, saliency detection , classification , subject extraction , gesture segmentation , economic load dispatch , shape design , and big data [66].

2.8.1 Ant Colony Optimization Algorithm

Hyper-heuristics are a collection of approaches aimed at automating the design and adjustment of heuristic methods to address difficult computational search problems. Their main goal is to develop more widely applicable search algorithms [67].

In 1992, Dorigo proposed the ant colony optimization (ACO) algorithm. It is an evolutionary algorithm based on heuristics. There are several iterations in ACO's algorithm. In each iteration, several ants use heuristic information and the collected experiences of others to build complete solutions. Ant colonies that once existed. These accumulated tales are Pheromone trail, which is deposited by the animal, is used to represent the animal on the components of a solution. In some cases, the pheromone may be the components and/or connections of a system solution that is specific to the problem at hand. The method by which here are some examples of the pheromone update rule [68].

The Operations Research community has extensively researched the Ant Colony Optimization (ACO) techniques for combinatorial optimization problems [69].

Ant System (AS) was the first algorithm to use the well-known traveling salesperson problem as an example application (TSP)A promising start was not enough to overcome the superiority of existing algorithms. However, it had a significant impact on the development of algorithmic research. On a wide range of applications, variants with significantly improved computational performance of various issues. There are, in fact, several uses for this technology.

Algorithms that achieve world-class results. An illustration of this would be the use of ACO algorithms. Solutions to issues such as balancing the assembly line, sequential ordering, and scheduling Two-dimensional high-performance protein folding DNA sequencing, ligand docking, and probabilistic TSP packet-switched routing in Internet-like networks, and so forth.

Assumption-Case Optimization gives existing applications and algorithmic variations a common framework. ACO algorithms are those that use the ACO metaheuristic as a guiding principle [70].

Marco Dorigo proposed and published the first ACO algorithm Ant System (AS), a set of three algorithms, as part of his doctoral thesis. After a few more years, the three technical reports on the algorithm were first published in the IEE Transactions on Electrical Engineering Humans, Systems, and Cybernetics. Among the three algorithms, the biggest difference was that pheromone was updated immediately after ants in terms of quantity and density transit cities while simultaneously depositing pheromones A function of tour quality was updated once all ants had completed building the path.

For an update to the pheromone level. Ant-improved cycle's performance was the catalyst for this development. Ant was based on research into two other algorithms: stop and ant-cycle. System. AS's first iteration had promising results, but they were not enough. to compete with well-known algorithms However, these results were disappointing. encouraging enough to arouse interest in further investigation in this area [71].

It is possible to pick a subset of data depending on one or more features. To reduce the size of a large dataset, this technique can be employed. To remove extraneous data from the data source and provide accurate predictions or outputs in analytics of huge data. Big data analysis is hampered by the generous size of the dataset.

The process of reducing a large data set is critical to effective analytics and prediction. It is essential in big data analytics to choose features that can be utilized to reduce a large data collection into smaller subsets [72].

2.8.2 Ant Colony Optimization Algorithm Work

The ACO algorithm is working in four stages: [71]

a. All parameters and pheromones variables are set up in the initialization step.

b. A group of ants build a solution to the problem at hand using the pheromone values and other information they have access to.

c. In this stage, ants can be used to improve on the solution that was already built.

d. Updates to pheromone variables based on ant search experience are included in the final stage of the global pheromone update.

2.8.3 GENETIC ALGORITHM

One of the earliest population-based stochastic algorithms to be proposed was the Genetic Algorithm (GA). Selection, crossover, and mutation are the primary GA operators, as they are in other algorithms. This chapter provides an overview of the algorithm and examines its performance in several real-world scenarios [73]. As a well-known metaheuristic algorithm, the Genetic Algorithm (GA) is based on the process of biological evolution. As in nature, GA mimics Darwin's theory of evolution by means of selection for fitness. In 1992, J.H. Holland produced the idea of GA. The most fundamental aspects of GA are chromosome representation, fitness selection, and biologically inspired operators. Inversion, a new element introduced by Holland, is commonly employed in GA implementations. The chromosomes are usually encoded as binary strings. There are two alleles (gene variants) for each chromosome locus, 0 and 1, in the human genome. Chromatin is thought of as a solution space point. Iteratively repopulating its population is used to do genetic operations. The population's fitness function is employed to assign values to each chromosome. Operators based on biological principles include selection, mutation, and crossover. During selection, chromosomes are chosen for further processing based on their fitness value. In the crossover operator, a random locus is selected, and the chromosomes are rearranged to make offspring. Some chromosome bits will be randomly flipped in mutation because of likelihood. Reliance on operators, representation, and fitness for the ongoing growth of GA has decreased [74].

2.8.4 Practical Swarm Optimization

A subset of the artificial intelligence (AI) area, swarm intelligence algorithms are becoming increasingly popular in addressing various optimization problems and have been widely used in a variety of applications. Numerous swarm intelligence algorithms have been created during the last few decades, including ACO, AFS, BFO, and (ABC) [75]. Autonomy is one of the most obvious advantages of SI. Without any external control, each agent in the swarm is responsible for their own actions and decisions. To put it another way, the agent here offers a potential solution to a certain problem. The second benefit is self-organization, which we can infer from that. Rather than focusing on a single agent, intelligence arises from the swarm itself [76]. As a result of its simple framework and fast convergence time, the Particle Swarm Optimization (PSO) technique has been successfully applied to numerous complex optimization problems [77] When using Particle Swarm Optimization, a group of particles moves together to discover the best possible solution.

To update the position of these particles, known as velocity vectors, a vector is used. The success rate of Particle Swarm Optimization can be improved in numerous ways. Analysis of real-world data and social models is at the heart of Particle Swarm Optimization (PSO). Particle Swarm Optimization is part of the Swarm Intelligence family; therefore, swarms or neurons work together to discover the best answer. [78].

A particle's exact location and velocity are of no consequence; instead, the swarm (or neuron) is given an initial start by being flung across hyperspace at random positions and speeds.

2.9 Feature Selection

It is mutual practice in machine learning to use feature selection (FS), especially when dealing with large, high-dimensional datasets. Input mining, Pattern Recognition, Image processing, and other Machine Learning algorithms demand less and less high-quality data as the amount of data grows exponentially. The "Curse of Dimensionality" is what Bellman refers to this situation as. Noise, irrelevant, and redundant data are more prevalent in higher-dimensional datasets. The overfitting of the model results in an increase in the learning algorithm's error rate. Techniques such as "Dimensionality Reduction" are used in the preprocessing stage to address these issues. As a result, dimensionality reduction techniques are most typically used in Feature Selection (FS) and Feature Extraction (FE).It is possible to choose a subset of features in FS based on the redundancy and significance of features. based on the qualities that are relevant and those that are superfluous [79]. Data pre-processing techniques such as feature selection have become an essential part of the machine learning process. In machine learning and statistics, this is also known as variable selection, attribute selection, or variable subset selection. Relevant features are identified, and redundant or noisy data is filtered out. This method speeds up data mining algorithms, enhances forecast accuracy, and makes data more understandable. A feature is considered irrelevant if it does not provide any valuable information, and redundant if it does not supply any additional information. When it comes to inductive learning, a supervised approach is preferable [80].

2.9.1 Feature Selection Types

The FS have three types deepened on contact with a learning model [79] .

A. Filter Feature Selection: - In this method, the model starts with all features and selects the best features subset based on statistical measures such as Pearson's correlation, Linear Discriminant Analysis (LDA), ANOVA, Chi-square, Wilcoxon Mann Whitney test, and Mutual Information (MI). All these statistical methods depend on the response and feature variable present in the dataset

B. Wrapper Feature Selection: - The classifier affects the Wrapper method's performance. Based on the results of the classifier, the best collection of characteristics is chosen. Wrapper methods are more computationally intensive than filter methods because of the number of learning steps and cross-validation that must be performed. In comparison to the filter method, these approaches are more precise. Recursive feature elimination, Sequential feature selection algorithms, and Genetic algorithms are only a few examples.

C. Embedded Method Feature Selection: - It selects features by a combination of ensemble and hybrid learning. As a result of its collective decision-making, it outperforms the other two models. One example of this is the random forest. Compared to wrapper approaches, this is a lot less resource intensive. A disadvantage of this approach is that it is based on a learning model.

2.9.2 Techniques for Finding Feature

The feature selections have some methods to find the features (Wan 2019)

A. Selecting one by one: The goal of FSS is to identify the most important qualities to include in a subset, while ignoring the less important ones. Every iteration, the greatest feature is found and added to an already-existing list of the top features. There is no further refinement after the adding of any additional features, therefore the search ends and the most important set of features is returned.

B. Selecting backwards in order (BSS): When it comes to determining which features are most important, BSS begins by looking at all of them and then seeks to eliminate the ones that are unnecessary or repetitive. Every iteration of the whole dataset is used to look for the feature that should be deleted. After a certain validation method, the next batch is analyzed. If a new feature subset's evaluation rate is higher than the prior subset's, the current best feature subset will be replaced. Every feature in the dataset is deleted until it reaches an empty set.

C. Climbing a mountain (HC): In HC, only one feature can be added or removed at a time. After searching through a random collection of features, it identifies the best features and then toggles each feature's current condition. When selecting the optimal set, the number of iterations is used to determine the stopping condition. When the last iteration reaches its limit, it returns the final collection of optimal features.

2.9.3 Feature Selection in Smart Campus

Smart campus has data sets that include the (temperature, lighting, humidity, noise, and air pressure). At the beginning of the process, the data set is separated into two parts: training and testing. It's then time for feature selection and feature scaling, which aims to choose the most crucial aspects that will have the greatest impact on model accuracy [81]. To improve the robustness and reproducibility of the classifiers, it was decided to train them on a smaller feature space instead of the original, larger feature space. Stability of a feature selection technique results in a consistent feature subset, regardless of the addition or removal of training data. Only when the training data is varied does a feature selection method remain stable [82].

2.10 PREVIOUS WORK

(Kim et al., 2018) Used real-time sensors and machine intelligence to improve learning experiences and easy communication between students and the teachers. Their system gave real-time suggestions improve the goodness and the ability of remembering of their offer by letting the offer to real-time adjustments/correcting to nonverbal communication for example gestures, facial expressions, and human body language.

(Aguilar et al., 2015) Used the multiagent systems concept, described a smart classroom in terms of these features. They defined these components as agents. They designed two agent frameworks that characterized the various sorts of components (software and hardware) in a smart schoolroom.

[83] recommended that learning analytics tasks are be defined as services that may be called by smart classroom components. They showed how to merge in a smart classroom the cloud and multi-agent paradigms to serve intelligent and non-intelligent entities with academic services, to enhance the learning process, they defined a smart classroom that can use a collection of learning analytics activities specified as services. [84] developed based virtual campus aide a Deep Neural Network (DNN). As a result of this work, the robot dialogue system now features Chinese Word Embedding, which significantly improves dialogue tolerance and semantic interpretation. Tokenize the Chinese phrase using the standard method of emotion identification.(Osegi, Anireh, and Taylor n.d.2020) used Auditory Machine Intelligence (AMI) approach is to forecast classroom tendencies and patterns in the weather using a test smart bed setup. They used thermal sensors and actuator signals, the AMI detects, predicts, and regulates the temperature in a classroom.

[86] they compared multiple IoT sensing technologies for class occupancy measurement in regarding how much it will set you back, how secure your personal data will be, how simple it is to use, and how accurate, and anticipate classroom attendance using AI approaches. Using them on real-world data and accurately projecting the future With an RMSE inaccuracy of 0.16, attendance is possible. [87] created four different experimental datasets to solve a problem with dataset imbalance, so they created them and used five classification algorithms to use multiple datasets to educate a classification model in order to solve it. These algorithms were Random Forest, J48, Nave Bayes, SMO, and Logistic regression to solve the problem of dataset imbalance.

[88] prepared ARAS dataset using contain primary, univariate, and correlation matrices. Data was analyzed using Logistic Regression (LR), SVM, and KNN algorithms to determine the accuracy of selected features. [89] proposed for HRES an internet of things (IoT) architecture, consisting of four components: wind turbine, photovoltaics, and battery

storage. Layers of power, data acquisition, system of transmissions, and application are all part of the proposed architecture.

[90] to increase the accuracy of activity detection in smart homes, they amended the data preprocessing and recognition phases, and more crucially, they presented a unique sensor segmentation approach and an update of KNN algorithm. The segmentation method divided sensor input into fragments based on predetermined activity information, and the suggested modified KNN algorithm classified the fragments using center distances. [91], proposed an innovative smart home system that used an algorithm (Support Vector Machine) for intellect decision making and used technicality based on the blockchain for making sure device identification and verification. Foundation of Raspberry Pi, 5 V relay, and other sensors are the components of their system. [92] proposed the use of Convolutional Neural Networks (CNNs) architecture as a deep learning model for recognizing human action [93] in their research, had developed a multi-agent control system with stochastic intelligent optimization, algorithms for multi-objective genetic programming (MOGA) and hybrid multi-objective genetic programming (HMOGA) in smart building.

[94] to maximize user comfort while minimizing energy consumption, they have developed an improved optimization function. A detailed analysis is carried out to come up with a comprehensive formulation for energy optimization, they have used genetic algorithm (GA) and particle-swarm optimization (PSO) algorithms to optimize [95] in their work addressed Quality of service (QoS)-based service composition issues using two metaheuristic algorithms: Genetic Algorithm (GA) and Particle Swarm Optimization (PSO).

3. METHODOLOGY

3.1 INTRODUCTION

In this chapter the way that is used to get the best results with high accuracy will be explained with details. Firstly, starting with describing the framework that we used in this thesis to oversee the problem of data by using feature selection in smart classroom environments. Figure below depicts the proposed framework whereas the details are given in subsequent sections.

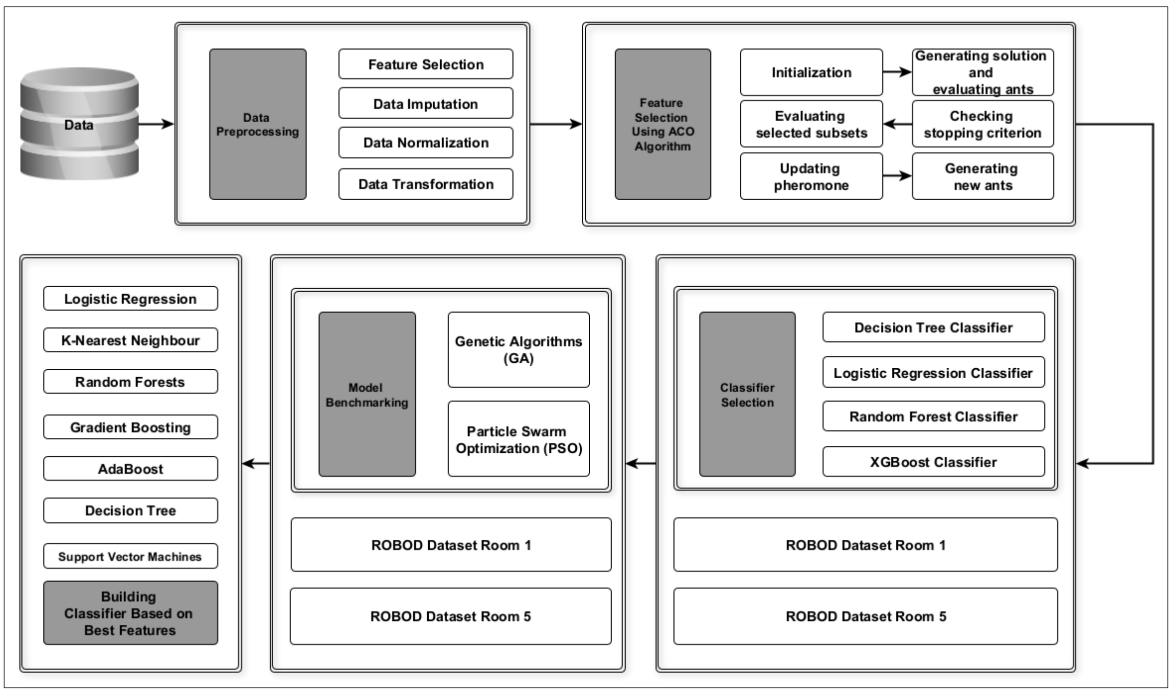

Figure 3.1: Proposed Structure

The details of the proposed framework are as follows:

a. we start the model with describing the datasets that we used in our experiments

b. data pre-processing which includes feature selection, data imputation, data normalization, and data transformation

c. modeling the Ant Colony Optimization for feature selection

d. choosing the best machine learning classifier to work with the ACO algorithm with experimentation on two datasets

e. model benchmarking using two metaheuristic algorithms with experimentation on two datasets

f. finally, building a classification model using the set of best features with experimentation on seven different machine learning classifiers.

3.2 DATASET DESCRIPTION

In our experiments, we focused on supervised learning approaches in which sensor data along with its features are fed into the system for processing.

Labeled data are required for both training the model and assessing its performance to assess the effectiveness of the proposed method. Our focus is on selecting only the features that are expected to be more relevant than the other features in our calculations. The dataset used in our experiments is titled as “Room-level Occupancy and Building Operation” or (ROBOD).

It is a comprehensive dataset with information related to outdoor weather conditions, HVAC operations, indoor environmental conditions, energy consumption (for example, HVAC, fans, plug loads and lighting), and Wi-Fi connected devices. The data was collected with the support of a set of heterogenous sensors from five rooms located in a university environment (the School of Design and Environment 4, building located at the National University of Singapore).

The dataset also comes with a ground truth occupant presence and count information for each of the five rooms. The room are not the same in size and they serve different purposes. For example, Room 1 and 2 are two different-sized lecture rooms, Room 3 is an office space for administrative staff, Room 4 is an office space for researchers, and finally Room 5 is a library space that is accessible to all students.

Data collection was not a smooth process as it took about 181 days to finish, where the data was collected from the five rooms at a sampling resolution of 5 minutes. The dataset can serve different purposes such as occupancy prediction, occupant behavior modeling, building simulation, building control, energy forecasting, and finally building analytics. Look at the description of each room in the following table.

Table 3.1: Room Description.

Room	Usage	Occupant type	Level	Floor Area (m2)	Floor to ceiling height(m)	Room volume (m3)	Seating capacity (person)	Maximum occupancy density (m2/person)
Room1	Lecture room	Students	4	118.6	4.1	486.2	40	3.0
Room2	Lecture room	Students	4	53.7	4.1	220.2	40	1.3
Room3	Office space	Admins staff	5	98.4	4.2	413.2	15	6.6
Room4	Office space	Researchers	3	141.9	4.1	581.7	25	5.6
Room5	Library space	Students	2	182.8	7.5	1363.3	36	5.0

Now, let us look at the summary statistics of ROBOD dataset:

Table 3.2: Summary statistics of ROBOD dataset

Item	Room	# Records	# Features
1	Room 1	8338	28
2	Room 2	8322	28
3	Room 3	8339	36
4	Room 4	13523	36
5	Room 5	13536	36

To be noted here that in our calculations, we picked up only two rooms for experimentation: Room 1 and Room 5. The reason is that they differ in both number of records and number of features. The features used for Room 1 and Room 5 were as follows:

a. Room 1: timestamp, VOC, sound pressure level, indoor relative humidity, air temperature, illuminance, pm2.5, indoor_CO2, WIFI connected devices, ceiling fan energy, lighting energy, plug load energy, chilled water energy, FCU fan energy, temp set point, FCU fan speed, supply air pressure, supply air temperature, biromantic pressure, dry bulb temp, global horizontal solar radiation, wind direction, windspeed, outdoor CO2, rainfall raw, outdoor relative humidity, occupant presence, occupant count.

b. Room 5: timestamp, VOC, sound pressure level, indoor relative humidity, air temperature, illuminance, pm2.5, indoor_co2, WIFI connected devices, ceiling fan energy, lighting energy, plug load energy, chilled water energy, AHU fan energy, supply airflow, damper position, temp set point, cooling coil valve position, cooling oil valve command, AHU fan speed, off coil airtime, off coil temp set point, pressure across filter, supply air humidity, supply air pressure, supply air temperature, biromantic pressure, dry bulb temp, global horizontal solar radiation, wind direction, windspeed, outdoor_CO2, rainfall raw, outdoor relative humidity, occupant presence, occupant count.

Look at the types of data gathered by sensors and the variables they measured [https://github.com/zeynepduygutekler/robod].

Table 3.3: Types of Data Gathered by Sensors

Indoor environmental quality	**Wi-Fi**	**Energy**	**HVAC Operations**	**Outdoor Weather** [c]	**Occupancy**
VOC	Wi-Fi connected Devices	Ceiling fan Energy	Supply air flow	Barometric pressure	Occupant presence
Sound pressure level		Lighting Energy	Temperature setpoint	Dry bulb temperature	Occupant count
Relative humidity		Plug load Energy	Cooling coil valve position	Global horizontal solar radiation	
Air temperature		Chilled water Energy	Cooling coil valve command	Wind direction	
Illuminance		AHU/FCU fan energy	AHU/FCU fan speed	Wind speed	
PM2.5			Off coil air temperature	CO_2	
CO2			Off coil temperature setpoint	Rainfall raw	
			Supply air humidity	Relative humidity	
			Pressure across filter		
			Supply air static pressure		

3.3 DATA PREPROCESSING

Real world data is usually messy and is typically created, processed, and stored by humans, applications, or business processes. As such, data may include missing fields, erroneous values, duplicate records, or tables. While humans could identify, locate, and manage such problems in data, machine learning or data mining applications do not have such a skill and therefore the data needs to be automatically preprocessed. Being part of data preparation, data preprocessing defines any kind of preprocessing applied to raw data to prepare it for other data processing operations. It includes steps that we need to follow to transform or encode the data so that it is easily parsed by the machine. One of the main objectives of data preprocessing is to transform data into a form that can easily and more effectively be processed in machine learning, data mining, or any other data-centric endeavors. Such techniques are applied in the preliminary stages of the machine learning pipeline to make sure that exact findings are achieved. A general view of the preprocessing steps can be found in Figure below:

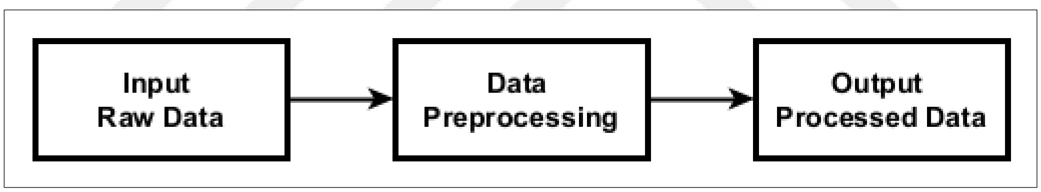

Figure 3.2: General View Of Data Preprocessing.

Several techniques and tools are associated with data preprocessing such as data sampling, transformation, denoising, feature selection and imputation. In the current study, we applied the following preprocessing steps:

A. Imputation: Replacing missing data with substituted values (e.g., 0, mean, max, min, etc.). Having missing data in the corpus can cause several issues such as introducing a substantial amount of bias, making the process of data analysis more complicated, and reducing efficiency. In our study, we applied data imputation by replacing missing values with zero.

B. Normalization: Organizing the data to appear similar across all the fields and records. It has the benefit of increasing the cohesion of data which would lead to have a higher quality data. When data normalization is done, the resulting data will be in a standardized information entry. In our study, we applied data normalization in the form of min-max normalization (in which zero represents no value and one is the maximum value).

C. Transformation: Converting raw data into a format (or structure) that is more suitable for subsequent operations in model building. It is an essential step in feature engineering as it facilitates discovering insights in data. Examples of data transformation includes converting categorical data into numeric, changing the distribution of the data to avoid biasness, data scaling, creating dummy variables, and so on. In the current study, we used data transformation to convert table data into a matrix format (i.e., NumPy array format) before using it in subsequent steps in model building

3.4 FEATURE SELECTION USING ACO ALGORITHM

FS is a process that selects a subset of salient features to build concrete learning models. Irrelevant or redundant features not only make the learning process more complicated, but also can diminish the efficiency of the learning model. By detecting and ignoring noisy and misleading features, the dataset quality improves. One way to discover irrelevant features is by checking if they are predictive of the decision feature or not. On the other hand, redundant features are those with high correlation to other features. Informative or useful features are those with high correlation to the predictor (label feature) and low correlation with other features.

Many of the swarm intelligence methods that mimic the social behavior of living beings (such as bees or ants) are being used in the field of feature selection. Among the many methods that were proposed for feature selection, population-based optimization methods like genetic algorithms and ant colony optimization that have attracted much attention recently. These methods try to achieve better results by application of knowledge from previous iterations.

3.5 ANT COLONY OPTIMIZATION ALGORITHM

Ant Colony Optimization (or ACO) algorithm belongs to a set of algorithms called “nature-inspired metaheuristic algorithms” that is gaining more popularity within the research field during the last few years.

The algorithm was motivated by the phenomenon that ants release pheromone on the ground on their way to and from the food source. The goal is to mark their preferable path so that other ants from the same colony can follow. The final goal is to find the shortest path between the colony and the food source. This is an excellent case of a self-organized process that uses positive feedback (i.e., releasing pheromone) and negative feedback (i.e., evaporation of pheromone).

Typically, optimization problems deal with minimizing or maximizing the computations of continuous n-dimensional functions through iteratively selecting n-inputs (within their bounds) and calculating the function value. Some of their uses include hyper-parameter optimization of machine learning algorithms, feature selection in machine learning, lost (or cost) function optimization for classical machine learning or deep learning problems, optimizing neural network architecture designs, best path finding.

3.5.1 Ant Colony Optimization for Feature Selection

ACO algorithm is a useful tool for selecting salient features and has been used in several studies. When the algorithm runs, many artificial ants iteratively traverse the feature space to construct the feature subsets.

The proposed algorithm is straightforward to implement with decreased computational complexity through using a machine learning classifier. We compare the performance of the proposed algorithm for maximizing efficacy of Genetic Algorithms and Particle Swarm Optimization on the task of feature selection using sensor data from a smart building environment. Results of calculations show an improvement in accuracy when using ACO algorithm as compared to the results of other metaheuristic algorithms.

Apart from diminishment in feature selection, accuracy in finding the subset of optimal features is vital. Therefore, in our work, feature selection is combined with a machine learning classification algorithm to learn and model the underlying processes.

The overall general process of ACO algorithm for feature selection is shown in the figure below. The procedure starts by generating a specific number of ants which are then put randomly on the graph (to put it another way, each ant has one random trait). Otherwise, the ant's number to put on the graph can be equal to the number of features within the data. After this, each ant begins constructing its own path at an unconventional character defining. Based on these initial positions, the ants traverse nodes probabilistically until the stopping criterion is met. The subsets that result from the previous phases are gathered and then evaluated. The algorithm may halt and output the best set of features encountered in case an optimal subset has been discovered or the algorithm execution times reach a certain limit.

It is also possible that the pheromone is updated, a new group of ants are generated, and the process starts over again if none of these halting conditions are met.

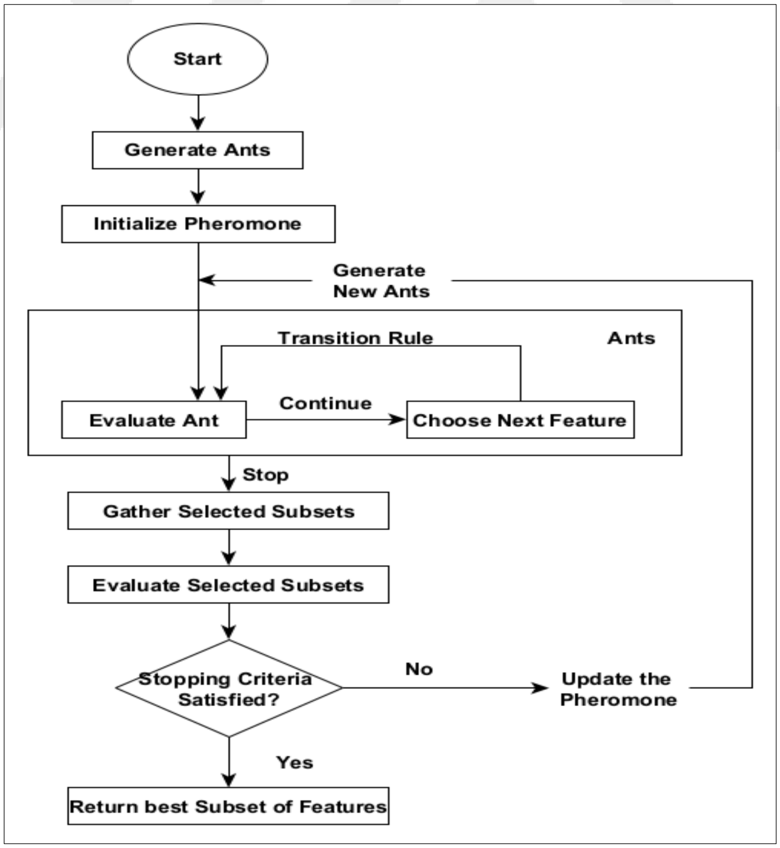

Figure 3.3: General ACO Processing Working For Feature Selection.

3.5.2 Proposed ACO Feature Selection Model

After we are done with data preprocessing, a typical process would include injecting the data into the model. However, we need first to apply feature selection to reduce data dimensionality. This process is shown in the figure below. What ACO algorithm does is that it explores the space of all subsets for the given features. The algorithm measures the effectiveness of the chosen feature subsets by invoking an evaluation function, along with using the resulting reduced feature space for classification evaluation. The recommended set of features to be used by the classification system is represented by the best feature subset found in the output.

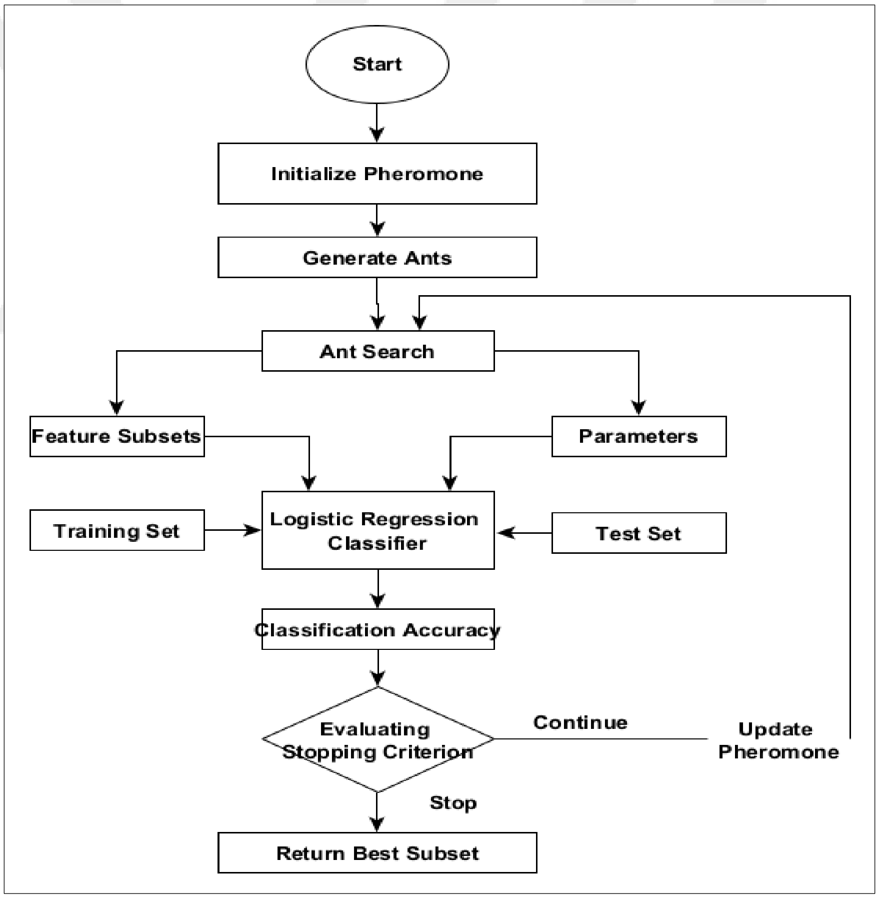

Figure 3.4: Proposed ACO Feature Selection Model

The main phases of the proposed feature selection model are as follows:

a. Initialization. This includes determining the population of ants, setting the density of pheromone trail connected to any feature, determining the maximum of allowed iterations.

b. Generating solution and evaluating ants. Assign ants randomly to features (exactly one ant to one feature) and visiting features where each ant would build a complete solution. The evaluation metric used in this phase is the Mean Square Error (or MSE) of the machine learning classifier. If any ant cannot reduce MSE of the classifier in successive phase, that ant will terminate its work and exit.

c. Evaluating selected subsets. This phase includes sorting subsets according to the classifier performance and their length. This is followed by selecting the best subset.

d. Checking stopping criterion. In this phase, if the number of iterations goes over the maximum allowed number of iterations, the process terminates, otherwise continue.

e. Updating pheromone. This phase includes reducing the pheromone concentrations of nodes followed by all ants depositing the quantity of pheromone on graph. As a last step, the best ant is allowed to deposit additional pheromone on nodes.

f. Generating new ants. This phase includes removing previous ants and generating new ones.

To be noted here that time complexity of the proposed model is O(lmn), such that *l* equals the number of iterations, *m* equals the number of ants, and *n* equals the number of original features. The worst-case scenario happens if each ant selects all the features. After each feature is added to the candidate subset, the heuristic is evaluated which results in *n* evaluations per ant. When the first iteration of the model is done, *mn* evaluations have been performed and after *l* number of iterations, the heuristic is evaluated *lmn* number of times.

3.6 CLASSIFIER SELECTION

The aim of this phase is to choose the best machine learning classifier to work with the ACO algorithm with experimentation on two datasets. To make use of the ACO algorithm's ability to perform hyperparameter search and determine which features are best for a given task, we integrate ACO algorithm in our solution to reduce the feature vector and increase the ‘classification accuracy.’

However, what criteria determine the viability of a proposed solution is the classification accuracy. The fitness function will return a number representing classifying how well each option works. The higher that number, the better the solution. To return the classification accuracy, we need a machine learning classifier that is trained on the feature elements returned by each solution. The fitness value that we get for each solution will allow us to select the best solutions as parents. These parents are then put in the mating pool to generate offspring which will be the elements of the new population of the new generation. For this purpose, we experiment the performance of ACO algorithm with four automatic classifiers based on machine learning: decision tree, logistic regression, random forest, and Boost. The experiments were done based on the settings shown below.

Table 3.4: ACO experimental settings.

I	# Experiments	# Ants	# Epochs	Alpha	Beta	Lambda	P_1	P_2
	5	20	40	1	0.2	0.005	0.75	0.1

The goal of ‘Alpha’ and ‘Beta’ parameters is to provide a relative importance of learning (i.e., pheromone) and heuristics. The goal of ‘lambda’ is to modify the objective function to minimize the number of selected features while maximizing the accuracy. In practice, higher values of ‘lambda’ would indicate more dominance of the number of characteristics of the aims. ‘P1’ and ‘P2’, (1-(p1+p2)) are the probabilities with which method #1, method #2, and method #3 of the heuristics are chosen The experimentations were done on two different datasets: Room 1 and Room 5. The reason is that they differ in both number of records and number of features which can be an indicator for the model performance across the different data sources. Summary statistics of these two rooms are as follows:

Table 3.5: Summary statistics of room 1 and 5.

Item	Room	# Records	# Features
1	Room 1	8338	28
2	Room 5	13536	36

The data was preprocessed by applying feature selection and dropping unnecessary features (i.e., columns) including ‘timestamp’ (the time when the observation was made) and ‘occupant count’ (number of occupants, if any). The other two features that we dropped were ‘tempest point’ and ‘off coil_ tempest point.’ The reason for dropping them is that their values were constant across the complete set of observations. Therefore, their existence was no longer necessary to the performance of the label feature. The label feature ‘occupant presence’ is the dataset ground truth and is set to ‘zero’ if no one is present in the room, otherwise it is set to ‘one.’

Besides feature selection, we also applied data imputation by replacing missing values with zero, data normalization in the form of min-max normalization (where 0 represents no value and 1 represents maximum value), and data transformation by converting table data into a matrix format (i.e., NumPy array format) because subsequent processes need the data to be in the matrix format.

3.7 MODEL BENCHMARKING

Model benchmarking usually refers to the evaluation and comparison of the proposed model against baseline model(s) regarding its ability to run as expected. It is an approach to identify the pros and cons of a given model in contrast to others. Comparisons could be made based on several metrics such as prediction accuracy, calculational difficulty, power to detect signal, model interpretability, and so on.

In the current study, we will benchmark our proposed technique (i.e., ACO algorithm and logistic regression classifier) for feature selection in smart classrooms environments against two other metaheuristic algorithms, namely Genetic Algorithms (GA) and Particle Swarm Optimization (PSO). For this purpose, we will use three machine learning classifiers: logistic regression, random forest, and decision tree

For genetic algorithms: number of generations were five, mutation rate was 0.8, and number of parents was two, while for the PSO algorithm the population size was ten. Dataset was split into training = 80% and testing = 20% of the entire dataset. The results of implementation on the two datasets: Room 1 and Room 5 are as follows (we picked the experiment with the highest result among all the five experiments).

3.8 BUILDING A CLASSIFIER USING BEST FEATURES

The previous phases shows that ACO algorithm with logistic regression classifier provides the best accuracy results. Accordingly, we decided to use the resulting features to build a machine learning classification model that can give predictions on new unseen data for smart classroom environments. For this purpose, we trained seven classifiers: support vector machines, decision trees, AdaBoost, Gradient Boosting, random forest, k-nearest neighbor, and logistic regression on Room 1 dataset.

Among the 25 features provided by sensors in Room 1, ACO algorithm with logistic regression classifier gave us the following 16 features: ceiling fan energy, FCU fan speed, illuminance, biromantic pressure, indoor CO2, WIFI connected devices, wind direction, plug load energy, lighting energy, VOC, global horizontal solar radiation, outdoor relative humidity, dry bulb temp, FCU fan energy, air temperature, and chilled water energy.

3.9 EVALUATION MEASURES

Typical evaluation measures used by researchers include confusion matrix, accuracy, precision, recall, specificity, F1-score, Precision-Recall (or PR curve), ROC (Receiver Operating Characteristics) curve, and PR vs ROC curve. In our study, the main evaluation metric was 'Accuracy.'

Accuracy in the machine learning context is the primary evaluation measure. It is a ratio of how many predictions were correct out of all the predictions. It can be calculated through the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.1}$$

where *TP* refers to true positive predictors, *TN* refers to true negative predictors, *FP* refers to false positive predictors, and False negative predictors, abbreviated *FN*, are a common type of statistical error. In addition to accuracy, we also applied few other techniques that were conducted during the testing phase. Models were evaluated during model building in two ways: k-fold cross validation and unseen testing data.

a. K-fold cross validation: for this purpose, we used 10-fold cross validation (i.e., k=10). Cross validation helps the model to generalize when the model is trained on different subsets of the data. During this process, the training and validation datasets are combined and randomized and then split into ten groups. Within the ten resulting subsets of data, one subset is used for training and the others for validation.

b. Unseen testing data: this technique, on the other hand, simulates real world scenario when the model is trained on the entire training and validation data, followed by assessing it on unseen new data. The goal of analyzing the mode's performance on unseen data is to assess the degree to which, during model training, overfitting occurs. Higher model performance on unseen data means that the model is dependable and can be safely applied in real-world scenarios.

4. RESULTS

The following table shows the accuracy results of applying the four machine learning algorithms on both Room 1 and Room 5 datasets. The goal is to find out which classifier works best with the ACO algorithm for feature selection (we selected the experiment with the highest result among all the five experiments).

Table 4.2: Performance classification performance of machine learning algorithms.

	Decision Tree	Logistic Regression	Random Forest	XG Boost
Room 1	0.848656	0.927438624	0.869371	0.854521764
Room 5	0.848547286	0.898268071	0.870120302	0.885339939

The table 4.1 shows the best results of the classifier (logistic regression) that merging with ACO algorithm by providing the high accuracy (%92, 89%) for provided data set, table 4.1, 4.2 shows the same results according to the parameters in the tables below.

Table 4.2: Confusion matrix values for room 1

Name of algorithm	Logistic Regression
TP	114
TN	246
FP	21
FN	11

Table 4.3: Confusion matrix values for room 5

Name of algorithm	Logistic Regression
TP	103
TN	206
FP	22
FN	14

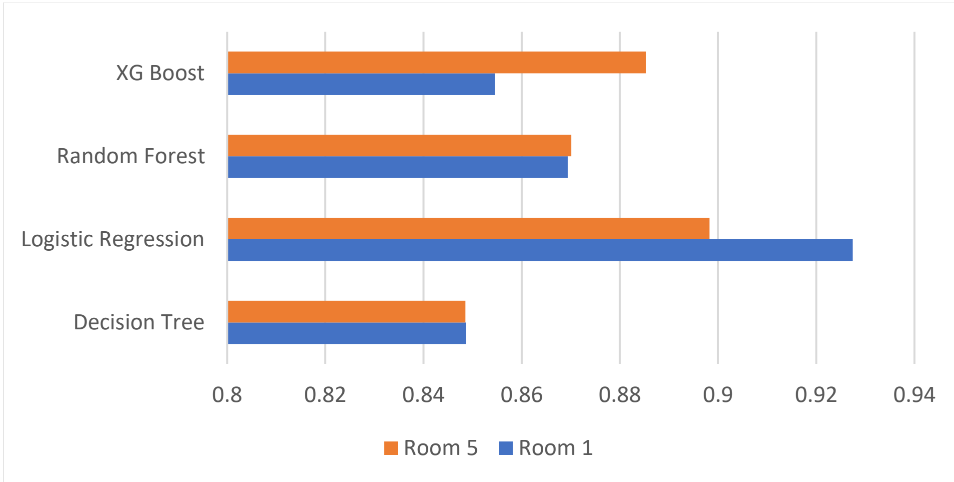

Figure 4.1: Performance Evaluation Of The Machine Learning Classifiers.

Building a classifier that using best features (which had been got by using logistic regression). The best results were founded by using random forest classifier were as shows in table and figure 4.4.

Table 4.4: Performance evaluation of classifiers on best features

Algorithm	**Logistic Regression**	**Random Forest**	**AdaBoost**	**Decision Tree**	**k-nearest neighbor**	**Gradient Boosting**	**Support Vector Machines**
Accuracy	0.939744	0.9784517	0.964884	0.967677	0.878292	0.974062	0.940143

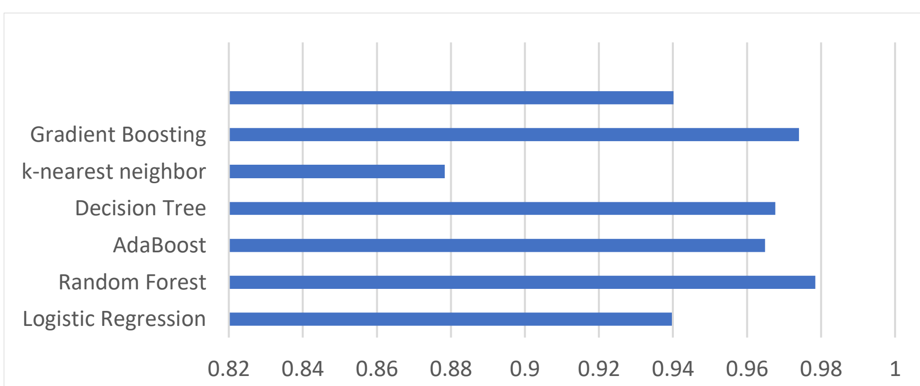

Figure 4.2: Performance Evaluation Of Classifiers On Best Features.

With the rapid advancement of technology, smart buildings, particularly smart classroom, have become increasingly important. For a smart campus to be successful, all the right conditions must be in place. As we concluded, the best features were created by combining the ACO algorithm with the logistic regression classifier. The random forest algorithm could take advantage of the new features to create a new model.

In this chapter we introduced the following results that we get it by making compering the performance (using the same data sets that collected from smart building) between our model (ACO and logistic regression classifier) with two metaheuristic (optimization algorithms) witch they

A. Genetic algorithm: - the using of feature selection with genetic algorithm shows the following results in table 4.3 and figure 4.3 which is less accuracy comparing with ACO algorithm.

Table 4.5: Performance evaluation of the genetic algorithm

	Random Forest	**Logistic regression**	**Decision tree**
Room 1	0.845601	0.790544	0.864153
Room 5	0.859675	0.865214	0.857090

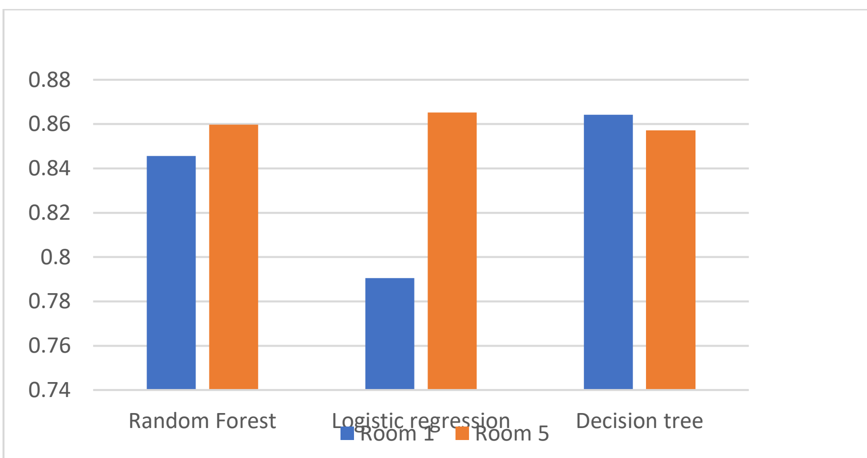

Figure 4.3: Performance Evaluation Of The Genetic Algorithm

B. Practical swarm optimization: - the using of PSO shows results that have accuracy less than using ACO with feature selection, table, and figure 4.4 shows these results

Table 4.6: Quantifying the particle swarm optimization algorithm's effectiveness.

	Random Forest	Logistic regression	Decision tree
Room 1	0.894343	0.861418	0.891349
Room 5	0.875161	0.872022	0.872022

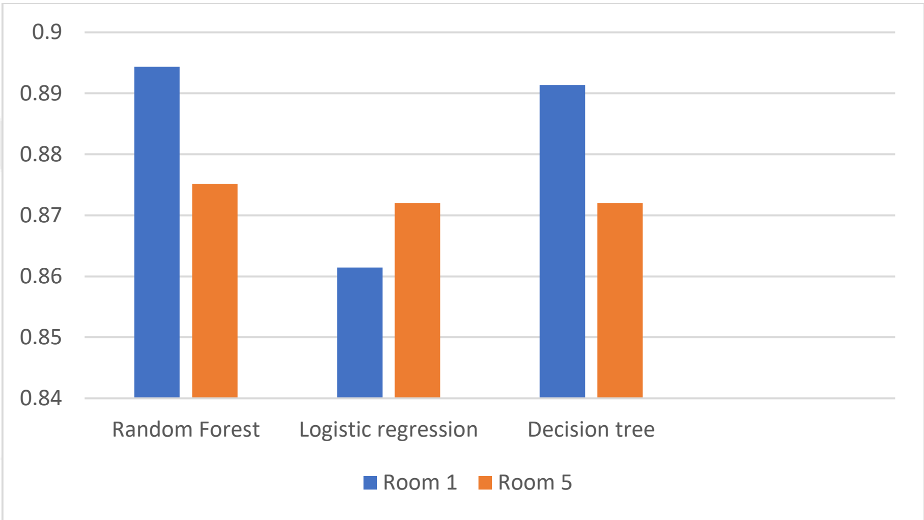

Figure 4.4: The Particle Swarm Optimization Algorithm's Effectiveness.

5. CONCLOUSIN AND FUTURE WORK

The using of feature selection which a machine learning operation in the choosing data sets from Room-level Occupancy and Building Operation" or (ROBOD) for the School of Design and Environment 4, building located at the National University of Singapore by taking the ant colony optimization algorithm with machine learning classifier (logistic regression) led to get the best features of (timestamp, VOC, sound pressure level, indoor relative humidity, air temperature, illuminance, pm2.5, indoor_CO2, WIFI connected devices, ceiling fan energy, lighting energy, plug load energy, chilled water energy, FCU fan energy, temp set point, FCU fan speed, supply air pressure, supply air temperature, biromantic pressure, dry bulb temp, global horizontal solar radiation, wind direction, windspeed, outdoor CO2, rainfall raw, outdoor relative humidity, occupant presence, occupant count) to use them for make classroom smart. After the compering between the results of (ant colony optimization algorithm, genetic, and SPO algorithms) which are belong to optimization algorithms and used in feature selection operations it had been founded that the accuracy of ACO is (92%) but the accuracy of genetic and SPO are (86% and 89%) respectively, so it is clearly that ACO algorithm had the best features when it used with machine learning classifier so it could be using

5.1 Future Work

After the conclusion of the high accuracy of using ant colony optimization which is merging with logistic regression classifier to get best features, a new model had been building by using these features with random forest classifier. This model could be using in the feature to make the operation of feature selection and applied it on any data sets of the buildings, the data set including (campus, homes, etc..) to make an appropriate smart building with appropriate conditions. The new model could easily determine the best features of (timestamp, VOC, sound pressure level, indoor relative humidity, air temperature, illuminance, pm2.5, indoor_CO2, WIFI connected devices, ceiling fan energy, lighting energy, plug load energy, chilled water energy, FCU fan energy, temp set point, FCU fan speed, supply air pressure, supply air temperature, biromantic pressure, dry bulb temp, global horizontal solar radiation, wind direction, windspeed, outdoor CO2, rainfall raw, outdoor relative humidity, occupant presence, occupant count) and other features of smart buildings in the future.

REFRENCE

[1] N. Iqbal, N. Ahmad, N. A. Basheer, and M. Nadeem, “Impact of corporate social responsibility on financial performance of corporations: Evidence from Pakistan,” International journal of learning and development, vol. 2, no. 6, pp. 107–118, 2012.

[2] H. Shi, B. Gao, and Z. Zhou, “A Survey on the Satisfaction of Diversified Learning Space in Colleges and Universities,” in Proceedings of the 6th International Conference on Education Reform and Modern Management (ERMM 2021), 2021, vol. 551, pp. 267–272. doi: 10.2991/assehr.k.210513.062.

[3] J. Faritha Banu, R. Revathi, M. Suganya, and N. R. Gladiss Merlin, “IoT based Cloud integrated smart classroom for smart and a sustainable campus,” Procedia Comput Sci, vol. 172, pp. 77–81, 2020, doi: 10.1016/j.procs.2020.05.012.

[4] Dick, S., “Artificial Intelligence”, Harvard Data Science Review, 1(1). 2019, doi:10.1162/99608f92.92fe150c.

[5] G. Carleo et al., “Machine learning and the physical sciences,” Rev Mod Phys, vol. 91, no. 4, p. 45002, 2019.

[6] M. Romero, W. Guédria, H. Panetto, and B. Barafort, “Towards a characterisation of smart systems: A systematic literature review,” Comput Ind, vol. 120, p. 103224, 2020.

[7] E. Alharbi, “People-Centric Smart Campus,” in 2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC), 2021, pp. 264–267.

[8] S. Hicks, C. Huang, E. Lee, K. Patel, and G. Vesonder, “Smart Campus for Better Study Spaces,” in 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2021, pp. 325–327.

[9] B. Liu, X. Li, Y. Wang, H. Wang, and F. Xu, “The System Framework of Data Mining and Learning Analysis for Smart Classroom,” in Proceedings - International Joint Conference on Information, Media and Engineering, ICIME 2018, 2019, pp. 331–336. doi: 10.1109/ICIME.2018.00077.

[10] S. F. Kuldeep, “IOT FOR SMART CLASSROOM IN IMPROVING TEACHING AND LEARNING APPROACH,” Asia-Africa Journal of Education Research, vol. 1, pp. 86–108, 2022.

[11] M. Sun, Z. Zhou, J. Wang, C. Du, and W. Gaaloul, “Energy-efficient IoT service composition for concurrent timed applications,” Future Generation Computer Systems, vol. 100, pp. 1017–1030, 2019.

[12] A. Zaballos, A. Briones, A. Massa, P. Centelles, and V. Caballero, “A smart campus’ digital twin for sustainable comfort monitoring,” Sustainability, vol. 12, no. 21, p. 9196, 2020.

[13] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, “A guide to machine learning for biologists,” Nat Rev Mol Cell Biol, vol. 23, no. 1, pp. 40–55, 2022.

[14] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, “Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-2019),” IEEE Access, vol. 9, pp. 26766–26791, 2021.

[15] Q. Al-Tashi, S. J. Abdulkadir, H. M. Rais, S. Mirjalili, and H. Alhussian, “Approaches to multi-objective feature selection: A systematic literature review,” IEEE Access, vol. 8, pp. 125076–125096, 2020.

[16] M. Al-Emran, S. I. Malik, and M. N. Al-Kabi, “A Survey of Internet of Things (IoT) in Education: Opportunities and Challenges,” in Studies in Computational Intelligence, vol. 846, Springer Verlag, 2020, pp. 197–209. doi: 10.1007/978-3-030-24513-9_12.

[17] S. C. Jia, R. D. Li, Y. J. Luo, and M. Zhu, “Smart Classroom System Based on Internet of Things Wireless Communication Technology,” Proceedings - 2021 International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2021, vol. 2021, pp. 154–157, 2021, doi: 10.1109/ICITBS53129.2021.00047.

[18] X. Wang, M. Li, and C. Li, “Smart classroom: Optimize and innovative-based on compared with traditional classroom,” International Journal of Information and

Education Technology, vol. 9, no. 10, pp. 741–745, 2019, doi: 10.18178/ijiet.2019.9.10.1296.

[19] A. N. Cormack, "See No…, Hear No…, Track No…: Ethics and the Intelligent Campus," Journal of Information Rights, Policy and Practice, vol. 3, no. 1, pp. 1–23, 2019, doi: 10.21039/irpandp.v3i1.59.

[20] L. S. Huang, J. Y. Su, and T. L. Pao, "A context aware Smart classroom architecture for smart campuses," Applied Sciences (Switzerland), vol. 9, no. 9, p. 1837, 2019, doi: 10.3390/app9091837.

[21] A. Solanki and A. Nayyar, "Green internet of things (G-IoT): ICT technologies, principles, applications, projects, and challenges," in Handbook of research on big data and the IoT, IGI Global, 2019, pp. 379–405.

[22] H. EL Mrabet and A. Ait Moussa, "Smart Classroom Environment Via IoT in Basic and Secondary Education," Transactions on Machine Learning and Artificial Intelligence, vol. 5, no. 4, 2017, doi: 10.14738/tmlai.54.3191.

[23] S. Madakam, "Internet of Things: Smart Things," International Journal of Future Computer and Communication, vol. 4, no. 4, pp. 250–253, 2015, doi: 10.7763/ijfcc. 2015.v4.395.

[24] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, and M. Mustaqim, "Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT Scenarios," IEEE Access, vol. 8, pp. 23022–23040, 2020, doi: 10.1109/ACCESS.2020.2970118.

[25] H. Aftab, K. Gilani, J. E. Lee, L. Nkenyereye, S. M. Jeong, and J. S. Song, "Analysis of identifiers in IoT platforms," Digital Communications and Networks, vol. 6, no. 3, pp. 333–340, 2020, doi: 10.1016/j.dcan.2019.05.003.

[26] H. Djelouat, A. Amira, and F. Bensaali, "Compressive sensing-based IoT applications: A review," Journal of Sensor and Actuator Networks, vol. 7, no. 4, p. 45, 2018.

[27] G. Cerullo, G. Mazzeo, G. Papale, B. Ragucci, and L. Sgaglione, "IoT and Sensor Networks Security," in Security and Resilience in Intelligent Data-Centric Systems

and Communication Networks, Elsevier, 2017, pp. 77–101. doi: 10.1016/B978-0-12-811373-8.00004-5.

[28] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Communications Surveys and Tutorials, vol. 17, no. 4, pp. 2347–2376, 2015, doi: 10.1109/COMST.2015.2444095.

[29] S. P. Raja and T. Sampradeepraj, "Internet of things: A research-oriented introductory," International Journal of Ad Hoc and Ubiquitous Computing, vol. 29, no. 1–2, pp. 4–14, 2018.

[30] Prof. Sathish and Dr. S. Smys, "A Survey on Internet of Things (IoT) based Smart Systems," Journal of ISMAC, vol. 2, no. 4, pp. 181–189, 2020, doi: 10.36548/jismac.2020.4.001.

[31] S. Bera, S. Misra, S. K. Roy, and M. S. Obaidat, "Soft-WSN: Software-defined WSN management system for IoT applications," IEEE Syst J, vol. 12, no. 3, pp. 2074–2081, 2018, doi: 10.1109/JSYST.2016.2615761.

[32] T. Kavitha and S. Silas, "A survey on energy harvesting routing protocol for WSN," in Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019, 2019, pp. 572–575. doi: 10.1109/ICOEI.2019.8862780.

[33] M. Elshrkawey, S. M. Elsherif, and M. Elsayed Wahed, "An Enhancement Approach for Reducing the Energy Consumption in Wireless Sensor Networks," Journal of King Saud University - Computer and Information Sciences, vol. 30, no. 2, pp. 259–267, 2018, doi: 10.1016/j.jksuci.2017.04.002.

[34] M. Ivanova, A. Boneva, and S. Ilchev, "Learning Performance Facilitation in a Sensor-Based Intelligent Classroom," in Big Data, Knowledge and Control Systems Engineering - Proceedings of the 7th International Conference, BdKCSE 2021, 2021, pp. 1–8. doi: 10.1109/BdKCSE53180.2021.9627308.

[35] A. S. Sadeq, R. Hassan, S. S. Al-Rawi, A. M. Jubair, and A. H. M. Aman, "A Qos Approach for Internet of Things (Iot) Environment Using Mqtt Protocol," in 2019

International Conference on Cybersecurity, ICoCSec 2019, 2019, pp. 59–63. doi: 10.1109/ICoCSec47621.2019.8971097.

[36] D. Dinculeană and X. Cheng, "Vulnerabilities and limitations of MQTT protocol used between IoT devices," Applied Sciences (Switzerland), vol. 9, no. 5, p. 848, 2019, doi: 10.3390/app9050848.

[37] D. Silva, L. I. Carvalho, J. Soares, and R. C. Sofia, "A performance analysis of internet of things networking protocols: Evaluating MQTT, CoAP, OPC UA," Applied Sciences (Switzerland), vol. 11, no. 11, p. 4879, 2021, doi: 10.3390/app11114879.

[38] B. Mishra and A. Kertesz, "The use of MQTT in M2M and IoT systems: A survey," IEEE Access, vol. 8, pp. 201071–201086, 2020, doi: 10.1109/ACCESS.2020.3035849.

[39] K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification," Augmented Human Research, vol. 5, no. 1, pp. 1–16, 2020, doi: 10.1007/s41133-020-00032-0.

[40] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," SN Comput Sci, vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.

[41] J. Wei et al., "Machine learning in materials science," InfoMat, vol. 1, no. 3, pp. 338–358, 2019, doi: 10.1002/inf2.12028.

[42] J. Díaz-Ramírez, "Machine Learning and Deep Learning," Ingeniare, vol. 29, no. 2, pp. 182–183, 2021, doi: 10.4067/S0718-33052021000200180.

[43] B. Abbasi and D. M. Goldenholz, "Machine learning applications in epilepsy," Epilepsia, vol. 60, no. 10, pp. 2037–2047, 2019, doi: 10.1111/epi.16333.

[44] D. Praveen Kumar, T. Amgoth, and C. S. R. Annavarapu, "Machine learning algorithms for wireless sensor networks: A survey," Information Fusion, vol. 49, pp. 1–25, 2019, doi: 10.1016/j.inffus.2018.09.013.

[45] S. Ray, “A Quick Review of Machine Learning Algorithms,” in Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Prespectives and Prospects, COMITCon 2019, 2019, pp. 35–39. doi: 10.1109/COMITCon.2019.8862451.

[46] A. Asselman, M. Khaldi, and S. Aammou, “Enhancing the prediction of student performance based on the machine learning XGBoost algorithm,” Interactive Learning Environments, pp. 1–20, 2021, doi: 10.1080/10494820.2021.1928235.

[47] S. C. Yu, X. Qi, Y. H. Hu, W. J. Zheng, Q. Q. Wang, and H. Y. Yao, “[Overview of multivariate regression model analysis and application].,” Zhonghua Yu Fang Yi Xue Za Zhi, vol. 53, no. 3, pp. 334–336, 2019, doi: 10.3760/cma.j.issn.0253-9624.2019.03.020.

[48] X. Zou, Y. Hu, Z. Tian, and K. Shen, “Logistic Regression Model Optimization and Case Analysis,” in Proceedings of IEEE 7th International Conference on Computer Science and Network Technology, ICCSNT 2019, 2019, pp. 135–139. doi: 10.1109/ICCSNT47585.2019.8962457.

[49] C. Bonte and F. Vercauteren, “Privacy-preserving logistic regression training,” BMC Med Genomics, vol. 11, no. 4, pp. 13–21, 2018.

[50] N. A. M. R. Senaviratna and T. M. J. A. Cooray, “Diagnosing Multicollinearity of Logistic Regression Model,” Asian Journal of Probability and Statistics, vol. 5, no. 2, pp. 1–9, 2019, doi: 10.9734/ajpas/2019/v5i230132.

[51] S. Sah, “Machine Learning: A Review of Learning Types,” ResearchGate, no. July, 2020.

[52] A. Tharwat, “Classification assessment methods,” Applied Computing and Informatics, vol. 17, no. 1, pp. 168–192, 2018, doi: 10.1016/j.aci.2018.08.003.

[53] B. Charbuty and A. Abdulazeez, “Classification Based on Decision Tree Algorithm for Machine Learning,” Journal of Applied Science and Technology Trends, vol. 2, no. 01, pp. 20–28, 2021, doi: 10.38094/jastt20165.

[54] D. Singh and B. Singh, “Investigating the impact of data normalization on classification performance,” Appl Soft Comput, vol. 97, p. 105524, 2020, doi: 10.1016/j.asoc.2019.105524.

[55] S. Qu, K. Li, S. Zhang, and Y. Wang, “Predicting Achievement of Students in Smart Campus,” IEEE Access, vol. 6, pp. 60264–60273, 2018, doi: 10.1109/ACCESS.2018.2875742.

[56] F. Y. Xu, Z. Q. Li, J. Q. Yue, and S. J. Qu, “A Systematic Review of Educational Data Mining,” Lecture Notes in Networks and Systems, vol. 284, pp. 764–780, 2021, doi: 10.1007/978-3-030-80126-7_54.

[57] P. Probst, M. N. Wright, and A. L. Boulesteix, “Hyperparameters and tuning strategies for random forest,” Wiley Interdiscip Rev Data Min Knowl Discov, vol. 9, no. 3, p. e1301, 2019, doi: 10.1002/widm.1301.

[58] A. B. Shaik and S. Srinivasan, “A brief survey on random forest ensembles in classification model,” in Lecture Notes in Networks and Systems, 2019, vol. 56, pp. 253–260. doi: 10.1007/978-981-13-2354-6_27.

[59] A. Ahmadian Ramaki, A. Rasoolzadegan, and A. Javan Jafari, “A systematic review on intrusion detection based on the Hidden Markov Model,” Stat Anal Data Min, vol. 11, no. 3, pp. 111–134, 2018, doi: 10.1002/sam.11377.

[60] T. Chadza, K. G. Kyriakopoulos, and S. Lambotharan, “Analysis of hidden Markov model learning algorithms for the detection and prediction of multi-stage network attacks,” Future Generation Computer Systems, vol. 108, pp. 636–649, 2020, doi: 10.1016/j.future.2020.03.014.

[61] I. Triguero, D. García-Gil, J. Maillo, J. Luengo, S. García, and F. Herrera, “Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data,” Wiley Interdiscip Rev Data Min Knowl Discov, vol. 9, no. 2, p. e1289, 2019, doi: 10.1002/widm.1289.

[62] G. Han, Y. Sun, Y. Feng, G. Lin, and N. Lu, “Machine Learning Regression Guided Thermoelectric Materials Discovery – A Review,” ES Materials & Manufacturing, vol. 14, pp. 20–35, 2021, doi: 10.30919/esmm5f451.

[63] D. Maulud and A. M. Abdulazeez, “A Review on Linear Regression Comprehensive in Machine Learning,” Journal of Applied Science and Technology Trends, vol. 1, no. 4, pp. 140–147, 2020, doi: 10.38094/jastt1457.

[64] S. Sun, Z. Cao, H. Zhu, and J. Zhao, “A Survey of Optimization Methods from a Machine Learning Perspective,” IEEE Trans Cybern, vol. 50, no. 8, pp. 3668–3681, 2020, doi: 10.1109/TCYB.2019.2950779.

[65] J. Stork, A. E. Eiben, and T. Bartz-Beielstein, “A new taxonomy of global optimization algorithms,” Nat Comput, pp. 1–24, 2020.

[66] W. Li, G.-G. Wang, and A. H. Gandomi, “A survey of learning-based intelligent optimization algorithms,” Archives of Computational Methods in Engineering, vol. 28, no. 5, pp. 3781–3799, 2021.

[67] F. M. Müller and I. S. Bonilha, “Hyper-heuristic based on aco and local search for dynamic optimization problems,” Algorithms, vol. 15, no. 1, p. 9, 2022, doi: 10.3390/a15010009.

[68] S. Liang, T. Jiao, W. Du, and S. Qu, “An improved ant colony optimization algorithm based on context for tourism route planning,” PLoS One, vol. 16, no. 9 September, pp. 20281–20292, 2021, doi: 10.1371/journal.pone.0257317.

[69] S. K. M. Hossain, S. A. Ema, and H. Sohn, “Rule-Based Classification Based on Ant Colony Optimization: A Comprehensive Review,” Applied Computational Intelligence and Soft Computing, vol. 2022, 2022, doi: 10.1155/2022/2232000.

[70] M. Dorigo and T. Stützle, “Ant colony optimization: Overview and recent advances,” International Series in Operations Research and Management Science, vol. 272, pp. 311–351, 2019, doi: 10.1007/978-3-319-91086-4_10.

[71] A. Akhtar, “Evolution of Ant Colony Optimization Algorithm -- A Brief Literature Review,” arXiv preprint arXiv:1908.08007, 2019.

[72] R. Joseph Manoj, M. D. Anto Praveena, and K. Vijayakumar, “An ACO–ANN based feature selection algorithm for big data,” Cluster Comput, vol. 22, no. 2, pp. 3953–3960, 2019, doi: 10.1007/s10586-018-2550-z.

[73] S. Mirjalili, “Genetic algorithm,” in Evolutionary algorithms and neural networks, Springer, 2019, pp. 43–55.

[74] S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: past, present, and future,” Multimed Tools Appl, vol. 80, no. 5, pp. 8091–8126, 2021.

[75] J. Tang, G. Liu, and Q. Pan, “A review on representative swarm intelligence algorithms for solving optimization problems: Applications and trends,” IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 10, pp. 1627–1643, 2021.

[76] L. Brezočnik, I. Fister Jr, and V. Podgorelec, “Swarm intelligence algorithms for feature selection: a review,” Applied Sciences, vol. 8, no. 9, p. 1521, 2018.

[77] F. Wang, H. Zhang, and A. Zhou, “A particle swarm optimization algorithm for mixed-variable optimization problems,” Swarm Evol Comput, vol. 60, p. 100808, 2021.

[78] L. Zajmi, F. Y. H. Ahmed, and A. A. Jaharadak, “Concepts, methods, and performances of particle swarm optimization, backpropagation, and neural networks,” Applied Computational Intelligence and Soft Computing, vol. 2018, 2018.

[79] B. Venkatesh and J. Anuradha, “A review of Feature Selection and its methods,” Cybernetics and Information Technologies, vol. 19, no. 1, pp. 3–26, 2019, doi: 10.2478/CAIT-2019-0001.

[80] V. Kumar and S. Minz, “Feature selection: a literature review,” SmartCR, vol. 4, no. 3, pp. 211–229, 2014.

[81] H. Singh, S. Gupta, C. Dhawan, and A. Mishra, “Path Loss Prediction in Smart Campus Environment: Machine Learning-based Approaches,” in IEEE Vehicular Technology Conference, 2020, vol. 2020-May, pp. 1–5. doi: 10.1109/VTC2020-Spring48590.2020.9129444.

[82] U. M. Khaire and R. Dhanalakshmi, “Stability of feature selection algorithm: A review,” Journal of King Saud University-Computer and Information Sciences, 2019.

[83] J. Aguilar, M. Sánchez, J. Cordero, P. Valdiviezo-Díaz, L. Barba-Guamán, and L. Chamba-Eras, “Learning analytics tasks as services in smart classrooms,” Univers Access Inf Soc, vol. 17, no. 4, pp. 693–709, Nov. 2018, doi: 10.1007/s10209-017-0525-0.

[84] P. S. Chiu, J. W. Chang, M. C. Lee, C. H. Chen, and D. S. Lee, “Enabling intelligent environment by the design of emotionally aware virtual assistant: A case of smart campus,” IEEE Access, vol. 8, pp. 62032–62041, 2020, doi: 10.1109/ACCESS.2020.2984383.

[85] E. N. Osegi, V. I. Anireh, and O. E. Taylor, “A Novel Neural Machine Intelligence Algorithm for Temperature Forecasting and Conditioning for Smart Classroom Applications”.

[86] T. Sutjarittham, H. Habibi Gharakheili, S. S. Kanhere, and V. Sivaraman, “Experiences with IoT and AI in a Smart Campus for Optimizing Classroom Usage,” IEEE Internet Things J, vol. 6, no. 5, pp. 7595–7607, 2019, doi: 10.1109/JIOT.2019.2902410.

[87] F. Wu et al., “Supporting poverty-stricken college students in smart campus,” Future Generation Computer Systems, vol. 111, pp. 599–616, 2020.

[88] J. W. Kasubi and D. H. Manjaiah, “Feature Selection Strategy for Multi-residents Behavior Analysis in Smart Home Environment,” in Data Management, Analytics and Innovation, Springer, 2022, pp. 11–26.

[89] A. M. Eltamaly, M. A. Alotaibi, A. I. Alolah, and M. A. Ahmed, “Iot-based hybrid renewable energy system for smart campus,” Sustainability, vol. 13, no. 15, p. 8555, 2021.

[90] Z. Wang, J. Gao, R. Chen, and J. Wang, “A modified KNN algorithm for activity recognition in smart home,” in Proceedings of the International Conference on Electronic Business (ICEB), 2018, vol. 2018-Decem, pp. 53–61.

[91] R. Majeed, N. A. Abdullah, I. Ashraf, Y. Bin Zikria, M. F. Mushtaq, and M. Umer, “An Intelligent, Secure, and Smart Home Automation System,” Sci Program, vol. 2020, 2020, doi: 10.1155/2020/4579291.

[92] H. D. Mehr and H. Polat, “Human Activity Recognition in Smart Home with Deep Learning Approach,” in 7th International Istanbul Smart Grids and Cities Congress and Fair, ICSG 2019 - Proceedings, 2019, pp. 149–153. doi: 10.1109/SGCF.2019.8782290.

[93] P. H. Shaikh, N. B. M. Nor, P. Nallagownden, I. Elamvazuthi, and T. Ibrahim, “Intelligent multi-objective control and management for smart energy efficient buildings,” International Journal of Electrical Power & Energy Systems, vol. 74, pp. 403–409, 2016.

[94] I. Ullah and D. Kim, “An improved optimization function for maximizing user comfort with minimum energy consumption in smart homes,” Energies (Basel), vol. 10, no. 11, p. 1818, 2017.

[95] N. Kashyap, A. C. Kumari, and R. Chhikara, “Service Composition in IoT using Genetic algorithm and Particle swarm optimization,” Open Computer Science, vol. 10, no. 1, pp. 56–64, 2020.