T.C.
ISTANBUL OKAN UNIVERSITY
INSTITUTE OF GRADUATE SCIENCES

THESIS
FOR THE DEGREE OF
MASTER OF SCIENCE
IN ADVANCED ELECTRONICS AND
COMMUNICATION TECHNOLOGIES PROGRAM

Hind Abduljaleel HAMEED

ENHANCING INDOOR POSITIONING PERFORMANCE THROUGH WI-FI/RSSI-BASED MACHINE LEARNING CLASSIFIERS

ADVISOR
Dr. Öğr. Üyesi Didem KIVANÇ TÜRELİ

ISTANBUL, January 2025

T.C.
ISTANBUL OKAN UNIVERSITY
INSTITUTE OF GRADUATE SCIENCES

THESIS
FOR THE DEGREE OF
MASTER OF SCIENCE
IN ADVANCED ELECTRONICS AND
COMMUNICATION TECHNOLOGIES PROGRAM

Hind Abduljaleel HAMEED
(203006012)

ENHANCING INDOOR POSITIONING PERFORMANCE THROUGH WI-FI/RSSI-BASED MACHINE LEARNING CLASSIFIERS

ADVISOR
Dr. Öğr. Üyesi Didem KIVANÇ TÜRELİ

ISTANBUL, January 2025

T.C.

ISTANBUL OKAN UNIVERSITY

INSTITUTE OF GRADUATE SCIENCES

YÜKSEK LİSANS TEZİ

ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Hind Abduljaleel HAMEED

(203006012)

ENHANCING INDOOR POSITIONING PERFORMANCE THROUGH WI-FI/RSSI-BASED MACHINE LEARNING CLASSIFIERS

Presentation Date of Thesis : January 10, 2025

Submission Date of Thesis : January 10, 2025

Advisor: Dr. Öğr. Üy. Didem KIVANÇ TÜRELİ__________________

Jury Members: Doç. Dr. Ömer Cihan KIVANÇ__________________________

Prof. Dr. Mehmet Serdar Ufuk TÜRELİ __________________

İSTANBUL, January 2025

To My Beloved Family

I express my deepest gratitude and appreciation for the unconditional support and unwavering assistance you have provided throughout the journey of preparing my master's thesis. You have been the true source of strength and motivation that pushed me to persist and overcome challenges.

Thanks to my dear mother, I cannot express how to thank her, without her I would not have reached this point nor reached what I am in my work.

A special dedication to the soul of my father, may God have mercy on him, who struggled in his life to leave us a solid foundation to support us against the shocks of life. Thank you, father, for what you did for us. It had a great impact on my personality and my determination to reach where I am now.

Thanks to my children who delayed me on this journey but were the biggest motivation to complete my journey and to be proud of me.

Also, thanks to my dear brother Mohanad and my dear sister Hadeel. You and your children have brightened my life with your unwavering moral support. Your presence in my life has played a major role in lifting my spirits and providing the support I needed in challenging times.

Thank you to everyone who taught me a letter in my life and had a great impact in shaping my personality. Thank you to Ms. Zahra, the math, language and science teacher who was the first to make me imagine the path to excellence in my life, and Ms. Sanaa, the English teacher who believed that I could excel more than I excelled, and all my professors at the University of Technology, Department of Control and Systems Engineering, and all my professors in the Master's program at Okan University, and special thanks to my dear supervisor, the esteemed Dr. Didem, who never hesitated to answer my questions and was my role model throughout my Master's journey.

With all my love and appreciation,

Hind.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to everyone who supported and guided me throughout the journey of completing this master's thesis. Without their continuous encouragement and assistance, this achievement would not have been possible.

TABLE OF CONTENTS

ABBREVIATIONS

5G, 6G	: 5 generation, 6 generation
AI	: Artificial intelligence
ANN	: Artificial neural network
AoA	: Angle-of-arrival
AP	: Access Point
Bluetooth	: Bluetooth is a short-range wireless technology
CNN	: Convolution neural network
EKF	: Extended Kalman Filter
ELM	: Extreme Learning Machine
FNNs	: Feed-Forward Neural Networks
GPS	: Global positioning system
IL	: Indoor Localization
IOT	: Internet of Things
IPS	: Indoor Positioning Systems
KNN	: K-nearest neighbours
LAN	: Local Area Network
LCD	: Liquid-crystal display
LDA	: Linear Discriminant Analysis
LOS	: Line-of-sight
LR	: Logistic Regression
LSTM	: Long-Short Term Memory
ML	: Machine learning
NLOS	: Non-line of sight

PCA : Principal Component Analysis

RFID : Radio Frequency Identification

RNN : Recurrent Neural Network

RSSI : Received Signal Strength Indicator

SVM : Support vector machines

TDoA : Time-difference-of-arrival

ToA : Time of arrival

UWB : Ultra-Wideband

WCL : Weighted Centroid Localization

SYMBOLS

Δt	The time difference.
c	The light speed.
θ	The arrival angle.
x, y	The coordination position.
d	The distance.
n	The signal propagation factor.
p_i	The probability of class i.
$K(x, x_i)$	The Kernel.
W_{in}	The input weight.
W_{out}	The output weight.

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

ENHANCING INDOOR POSITIONING PERFORMANCE THROUGH WI-FI/RSSI-BASED MACHINE LEARNING CLASSIFIERS

Indoor positioning (IP) is a pivotal component in real-time indoor localization (IL), contributing to the identification of user or device locations within confined spaces. Global positioning system (GPS) excels in outdoor positioning, but its efficacy diminishes indoors due to challenges like multipath propagation, non-line of sight, and signal distortion. To address this, an indoor positioning technique has developed, including Wi-Fi positioning system, five access points sensors, user as a tag, and based on the received signal strength power (RSS) with machine learning classifiers.

Machine learning empowers indoor positioning systems to adapt to diverse conditions, manage uncertainties, and enhance accuracy continually through data-driven learning. Fingerprinting localization faces challenges with high-dimensional data, addressed by dimensionality reduction methods like principal component analysis (PCA). Classification algorithms, such as Decision Trees (DT), Local Discriminator Analysis (LDA), Support Vector Machine (SVM), K-nearest neighbor (KNN), Logistic Regression (LR), Artificial Neural Networks (ANN), and Extreme Learning Machine (OS-ELM) are employed to extract key characteristics for localization and hence implement IP detection. The ensuing results are analyzed for accuracy, prediction speed, and training time.

The research sets a foundation for understanding the strengths and limitations of various classifiers in indoor positioning. The comparative analysis reveals that OS-

ELM exhibits exceptional accuracy, rapid prediction speed, and minimal training time, positioning it as a promising choice for real-time applications. The study concludes by outlining future research directions, emphasizing the refinement of OS-ELM and hybrid approaches to enhance accuracy and adaptability in dynamic indoor environments.

Keywords: Indoor Positioning (IP), Wi-Fi Positioning System, RSS Power, machine learning, dimensionality reduction, Extreme Learning Machine (ELM), training time, training and testing Accuracy.

ÖZET

WI-FI/RSSI TABANLI MAKINE ÖĞRENIMI SINIFLANDIRICILARIYLA KAPALI ALAN KONUMLANDIRMA PERFORMANSININ ARTIRILMASI

Kapalı alan konumlandırma (IP), gerçek zamanlı kapalı alan lokalizasyonunun (IL) önemli bir bileşenidir ve kullanıcı veya cihaz konumlarının kapalı alanlarda belirlenmesine katkıda bulunur. Küresel Konumlandırma Sistemi (GPS), açık alanlarda etkili bir şekilde çalışsa da, çoklu yol yayılımı, görüş hattı olmaması ve sinyal bozulması gibi zorluklar nedeniyle kapalı alanlarda etkinliği azalır. Bu sorunu çözmek için, Wi-Fi konumlandırma sistemi, beş erişim noktası sensörü, etiket olarak kullanıcı ve alınan sinyal gücü (RSS) ile makine öğrenimi sınıflandırıcılarına dayalı bir kapalı alan konumlandırma tekniği geliştirilmiştir.

Makine öğrenimi, kapalı alan konumlandırma sistemlerinin çeşitli koşullara uyum sağlamasını, belirsizlikleri yönetmesini ve veri odaklı öğrenme yoluyla doğruluğu sürekli artırmasını mümkün kılar. Parmak izi tabanlı lokalizasyon, yüksek boyutlu verilerle ilgili zorluklarla karşılaşır ve bu zorluklar temel bileşen analizi (PCA) gibi boyut azaltma yöntemleriyle ele alınır. Karar Ağaçları (DT), Yerel Ayırt Edici Analiz (LDA), Destek Vektör Makinesi (SVM), K-en yakın komşu (KNN), Lojistik Regresyon (LR), Yapay Sinir Ağları (ANN) ve Aşırı Öğrenme Makinesi (OS-ELM) gibi sınıflandırma algoritmaları, konumlandırma için temel özellikleri çıkarmak ve IP tespiti uygulamak için kullanılır. Ortaya çıkan sonuçlar, doğruluk, tahmin hızı ve eğitim süresi açısından analiz edilir.

Araştırma, kapalı alan konumlandırmada çeşitli sınıflandırıcıların güçlü ve zayıf yönlerini anlamak için bir temel oluşturur. Karşılaştırmalı analiz, OS-ELM'nin olağanüstü doğruluk, hızlı tahmin hızı ve minimal eğitim süresi ile öne çıktığını ortaya koyarak, onu gerçek zamanlı uygulamalar için umut verici bir seçenek haline getirmektedir. Çalışma, OS-ELM'nin ve dinamik kapalı alan ortamlarında doğruluk ve uyarlanabilirliği artırmak için hibrit yaklaşımların iyileştirilmesini vurgulayarak gelecekteki araştırma yönlerini ortaya koymaktadır.

Anahtar Kelimeler: Kapalı Alan Konumlandırma, Wi-Fi Konumlandırma Sistemi, RSS Gücü, makine öğrenimi, boyut azaltma, Aşırı Öğrenme Makinesi (ELM), eğitim süresi, eğitim ve test doğruluğu.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Indoor positioning (IP) is crucial for determining the location of a user or device within a confined space, constituting real-time indoor localization (IL). While the global positioning system (GPS) effectively serves outdoor positioning needs, its accuracy significantly diminishes in indoor environments due to challenges like multipath propagation, non-line of sight, and signal distortion. This limitation has led to the development of various indoor positioning techniques to address these issues [1]. The applications of Indoor Positioning Systems (IPS) technology extend across diverse sectors, including healthcare, building management, industry, police investigation, disaster management, and various other domains. Additionally, IPS technology finds relevance in emerging technologies like the Internet of Things (IoT) and smart architectures [2][3].

1.2 Challenges of Localization Techniques

In today's world, with the increasing use of smart devices, IP or IL has become a crucial task. It involves determining the location of a person or object within a building, which is different from determining their location outside. The indoor environment poses many challenges such as signal propagation leading to shadowing, scattering, attenuation, and distortion. These challenges have led to scientific research to discover new ways of indoor localization using various elements. However, the challenges are significant due to the variety of sizes and shapes of buildings, the presence of fixed and moving objects, as well as their influence on both line-of-sight (LOS) and non-line-of-sight (NLOS) radio blind spots, all have an impact on indoor positioning performance. Many solutions have been suggested for the recent years including time of arrival (TOA) [4], angle of arrival (AOA) [5], received signal strength (RSS) [6], difference of arrival (DOA) [7], time difference of arrival (TDOA) [8], and hybrid methods [9][10].

Indoor localization systems utilizing RSS do not require any specialized components, which makes their deployment much easier. These systems can be implemented through various technologies such as WiFİ [11], RFID [12][13], Bluetooth [16], and others [14][15][17].

RSS-based localization methods are generally categorized into two types: model-based (or range-based) methods [18] and fingerprinting-based methods [19]. Model-based approaches leverage estimate the distance using an equation which quantifies the relationship between received power and distance [20]. However, their performance may be affected by the complexities of indoor signal propagation environments.

Machine learning techniques, coupled with biometric approaches, are employed to match a user's position with a predefined set of positions. The average RSSI received at each position is retrieved from a database from data collected during system setup [21]. This data can be updated in time to match changes in the environment. Fingerprinting, which usually uses RSSI signals on WiFi has become very popular due to the wide deployment of WiFi. While techniques such as TDOA can give higher accuracy, RSSI is more robust to use in different environments without needing specialized hardware such as antenna arrays. It exhibits satisfactory performance in both LOS and NLOS environments.

The fingerprinting Wi-Fi localization occurs in two stages: the offline training stage and online usage stage. In the offline phase, a fingerprint database is established using Wi-Fi access points, which are then employed to estimate positions based on received RSSI signals during the online phase. These access points are collected from smart devices or signal transmitters and are tailored to the project's working environment in the form of radio maps.

However, the challenge of limited RSS datasets poses difficulties for effective machine learning models, which have shown promise in indoor-based localization applications. To overcome this challenge, machine learning systems have demonstrated the most promising performance due to their robustly and processing time [22][23]. Additionally, researchers have explored advanced deep learning techniques such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) for collecting and training RSS radio map databases [24][25].

1.3 The Problem Statement

Sensing data without knowledge of sensor locations renders a sensor network ineffective in its primary function of collecting and forwarding data. This underscores the critical need for precise sensor location determination in IP systems, particularly in challenging indoor environments with various obstacles. Existing approaches, including the use of ML to improve distance estimates between anchor nodes with known positions and a target with an unknown position or increasing the number of anchor nodes, are constrained by reliance on specific databases, limiting the adaptability of indoor positioning systems to dynamic environments.

Furthermore, various researchers have explored the development of reliable IPS using both single and hybrid machine learning algorithms. Despite these efforts, earlier mechanisms fell short in achieving a high detection positioning rate. Consequently, this thesis concentrates on attaining a high detection positioning accuracy for indoor positioning systems. The approach involves creating an integrated work mechanism that prioritizes IPS accuracy and processing time.

1.4 The Work Aims and Objectives

Our work aims to address key factors such as accuracy, complexity, reliability, and deployment in indoor positioning systems. The significance of precise distance estimates between anchor nodes and the mobile target is emphasized to achieve high-accuracy positioning. As a result, the aim of this work is to design an Indoor Positioning System (IPS) based on the RSS signals, Wi-Fi connections, with machine learning abilities that can specify characteristics involves addressing several key aspects to meet the outlined requirements:

Wireless Contact:

Utilize wireless technologies such as Wi-Fi, Bluetooth, for communication between the positioning system components. Wireless connectivity enhances flexibility and ease of use.

Easy for Personal Use:

Develop a user-friendly interface accessible via smartphones or dedicated devices. Implement intuitive features, minimal setup requirements, and straightforward calibration processes to ensure ease of use for individuals.

Low Cost:

Employ cost-effective hardware components and leverage existing infrastructure, such as Wi-Fi routers or Bluetooth beacons, to minimize expenses. Optimize the design for affordability without compromising on performance.

High Accuracy:

Implement advanced positioning algorithms, including machine learning techniques, to enhance accuracy. Utilize a combination of signal strength, time-of-flight, or trilateration methods to achieve precise indoor location tracking.

Suitable in All Indoor Environments:

Account for diverse indoor environments, considering factors like building layouts, obstructions, and signal interference. Employ a scalable system that adapts to different indoor settings, ensuring consistent performance across varied conditions.

1.5 The Main Contributions

To accomplish the objective of creating a real-time system capable of handling various scenarios and capturing data from a typical indoor environment. We've made the following issues:

1. Provide and calibrate a Wi-Fi-based indoor localization algorithms and implement them to determine suitable ML approaches.

2. Collect measurements to create a dataset within a single floor building to evaluate localization algorithms

3. Developing and applying different Machine Learning algorithms with offline and online behavior, as well as reporting the achieved reasonable accuracy.

4. Simulate the proposed IP systems used the MATLAB environment with no extra hardware or software.

5. Analyze and evaluate an appropriate option in which the user is simply carrying a cell phone. The phone does not have any special software loaded.

1.6 Thesis Organization

This thesis is organized as follows:

Chapter One: This section serves as the introduction and provides a clear explanation of the motive and objectives behind the selection of the thesis theme.

Chapter Two: Present a comprehensive description of the research background through providing a literature review.

Chapter Three: Focuses on detail the proposed indoor positioning system including the setup experiment and the database collection in different environments. Theories of all developed ML approaches are explained in details.

Chapter Four: Describes and discuss the experimental findings, provides the performance assessment of the detection rate for all recommended classifiers, and compares them to several detection strategies from prior research.

Chapter Five: Delivers conclusions, recommendations and suggestions for future works.

CHAPTER 2. THEORETICAL BACKGROUND AND LITERATURE REVIEW

2.1 Introduction

Localization techniques refer to the methods and technologies used to determine the precise location of an object or individual. Indoor localization techniques are specifically designed for determining the position of objects or individuals inside buildings, structures or enclosed spaces. On the other hand, outdoor localization techniques are used for determining the position of objects or individuals in open spaces, such as outdoors or in large areas. The main difference between indoor and outdoor localization techniques lies in the challenges posed by the different environments. Indoor environments present more obstacles, such as walls and furniture, which can interfere with signal transmission and accuracy. Indoor localization techniques include technologies such as Wi-Fi positioning, Bluetooth beacons, Infrared sensors, ultrasonic sensors, and radio frequency identification. Also, Ultra Wide Band (UWB) and their techniques with the various wireless IPs communication technologies are widely adopted. These techniques use signals and sensors to triangulate and estimate the location of an object or individual within an indoor space. Applications of indoor localization techniques include: asset tracking in warehouses, indoor navigation and wayfinding in large buildings such as airports or shopping malls, proximity marketing, location-based services for healthcare facilities such as tracking patients and medical equipment, indoor security systems, and personalized experiences in museums or exhibitions. Outdoor localization techniques, on the other hand, make use of technologies such as GPS, GNSS. This chapter, will provide a comprehensive description of the research background through a literature survey related to indoor positioning. The indoor localization system with their various methods tolerate for continuous tracking of the location of objects or humans are employed extensively. They are used as an efficient tools for measuring distances or/and angles. Determining the position of a tag is a process achieved in two primary steps. Initially, it entails acquiring a collection of accurate ranging data between the anchor and the tag. Subsequently, the obtained data undergoes processing to ascertain the probable location of the mobile station, commonly known as localization. Utilizing diverse machine learning approaches empowers indoor positioning systems to adapt to varying conditions, handle uncertainties, and continually improve accuracy through learning from data.

2.2 Literature Review

Many IP studies were investigated and discussed using different techniques and methods, among them that were related to Wifi with RSS and machine learning:

Hou et. al. (2015) [26] describe an indoor positioning system using lights. A LED lamp is used as the transmitter and a smartphone is the receiver. Positioning uses both the signal strength and angle of arrival. The outcomes revealed an average positioning error of around 10.2 cm within a (2×2×2.5) m indoor environment.

Kotaru et. al. (2015) [27] propose an indoor localization system named SpotFi for deployment on standard Wi-Fi. The localization algorithm was evaluated in a large indoor office space employing six Access Points. The median accuracy attained was 40 cm, showcasing resilience in the presence of indoor obstacles.

Zou et. al. (2015) [28] proposed a localization algorithm for indoor positioning utilizing an online sequential extreme learning machine (OS-ELM). Experimented within an Internet of Things Laboratory, the algorithm demonstrated an accuracy of 1.794 m within a testbed area of approximately 400 m2 (20 m × 20 m), featuring four Access Points (APs).

Thuong et. al. (2016) [29] investigate diverse facets of location fingerprinting-based indoor positioning systems, with a particular emphasis on factors influencing positioning accuracy, including the impact of human behavior on RSSI distribution. The experiments were carried out in an indoor space which was again within the school setting, covering an area of (7 m × 10 m). The results showcased positioning accuracy within the range of (2.0 - 2.5) m.

Ding et. al. (2017) [30] introduced a technique for fingerprint-based indoor positioning (IP) was proposed, known as the AP weighted multiple matching closest neighbor's approach. This method aims to overcome challenges observed in the conventional Weighted K-Nearest Neighbor (WKNN) algorithm for localization.

Dari and Pranowo (2018) [31] introduced a system combining RSS fingerprint method with KNN for enhanced location recognition in mobile-based wireless positioning systems utilizing wireless 802.11b technology.

Hou et. al. (2018) [32] proposed a wireless indoor localization approach, employing AoA and coordinates from Wi-Fi Access Points (APs) for outpatient wayfinding in hospitals. Despite its limited applicability in non-line of sight conditions, the system demonstrated a localization error of less than two and a half meters in 80% of the instances.

Oras et. al. (2019) [33] introduce a two-step localization algorithm, optimizing the deployment of APs based on RSS simulation measurements. The algorithm achieved an optimal location with a range probability of 0.74m and the highest RSS mean received power value of -45 dBm.

Abbas et. al. (2019) [34] introduced WiDeep, a deep learning-based indoor localization technique implemented on Android phones. Evaluated in university and residential settings, WiDeep consistently achieving an average precision in the 1-2m range on the two testbeds.

Xue et. al. (2020) [35] introduced the "Weight Range Localizer" and "Relative Span Exponential Weight Range Localizer" algorithms for RSSI localization in wireless sensor networks for short-range communication. The proposed algorithms demonstrated acceptable accuracy, achieving less than 1 m in IP.

Yan et. al. (2021) [36] proposed novel ELM localization technique for multi-floor environments, relying solely on RSSI fingerprints. The approach includes a specialized data preprocessing algorithm to efficiently handle extensive training and online measurement data in multi-floor settings. The offline phase utilizes individual ELMs for all floors, generating floor-level classification functions and position regression functions. In the online phase, first the algorithm determines a coarse localization step estimates the floor using floor-level classification functions, followed by a refined step for on-floor position estimation. Comparative experiments demonstrate significant performance advantages in both floor estimation and on-floor localization compared to existing algorithms.

Wang and Park (2021) [37] proposed a hybrid fingerprint location technology based on both RSS and Channel State Information (CSI). The methodology involves preprocessing RSSI and CSI values using Kalman filter and Gaussian function in the off-line phase, eliminating mutation and noisy data. The accurate hybrid fingerprint database is established, and the weighted k-nearest neighbor (WKNN) algorithm is employed for online positioning. Experimental results demonstrate the proposed algorithm is robust to noise, showcasing higher accuracy and smaller positioning errors.

Djosic et. al. (2022) [38] present a novel localization method to improve UWB accuracy with non-line-of-sight (NLOS) conditions. The approach uses multiple deterministic algorithms based on LOS-measured distances, which converge much more quickly to a solution and achieve better accuracy in LOS conditions, but also use fingerprinting-based algorithm for NLOS conditions. Experimental results show that this strategy outperforms traditional fingerprinting-based approaches.

Shyam et. al. (2023) [39] discuss utilizing UWB technology and a triangulation algorithm for monitoring and tracking assets in industrial settings. Operating in an edge computing environment, the system ensures quick and accurate data delivery to end-users. Wall-fixed anchors gather and process assets' location information, accessible to users through smart devices. The system traces assets in 2D and 3D, demonstrating increased accuracy and coverage. Experiments in a room and warehouse environment show improved performance, with special attention to error mitigation resulting in a significant reduction in error percentage.

2.3 Wireless Positioning Technologies

According to accuracy, stability, scalability, safety, complexity, and cost metrics criteria, several wireless positioning technologies can be created or employed to assess the efficacy of positioning systems [40]. There are many wireless positioning techniques:

2.3.1 Ultra-Wide-Band Technology (UWB)

Ultra-Wideband technology (UWB) refers to a telecommunication technology that utilises a frequency that is higher than 500MHz. It offers promising capabilities for accurate and reliable indoor positioning [14]. Its fine time resolution and ability to penetrate obstacles make it a suitable choice for indoor environments where GPS signals may be obstructed. Unlike other technologies such as Wi-Fi, Bluetooth, and RFID, UWB does not rely on signal strength or line-of-sight communication. Instead, it utilizes the transmission and reception of short radio pulses with sub-nanosecond durations. This enables UWB systems to provide centimeter-level accuracy in measuring distances, making them well-suited for applications that require high precision [19].

UWB technology offers various advantages and disadvantages, with different types and measurements. Advantages of UWB include high bit rate availability, low power consumption, low costs, and high accuracy positioning capabilities [41]. On the other hand, UWB technology also has its limitations. Some of the disadvantages of UWB technology include limited range, susceptibility to interference from other wireless devices, and regulatory restrictions on power levels. There are different types of UWB measurements used. These measurements include distance measurement, time-of-flight measurement (ToF), and signal strength measurement.

Additionally, there are different types of UWB antennas that have been proposed to meet the requirements of UWB technology [42]. Some examples of UWB antennas include planar antennas, printed antennas, and monopole antennas. These antennas are designed to provide the necessary bandwidth and frequency range for Ultra-Wideband communication.

2.3.2 Radio Frequency Identification (RFID)

Radio Frequency Identification (RFID) is a technology that stores and retrieves data via electromagnetic transmission to an RF-compatible integrated circuit, and it is gaining popularity for its role in enhancing data processing operations [13]. An RFID system is made up of many components, including RFID readers, which read data from RFID tags, and RFID tags themselves, as well as the communication between them. RFID tags may be passive or active. Passive RFID tags lack a battery, limiting their range to around (1-2) meters [43]. Active RFID tags, on the other hand, are fitted with small transceivers that can communicate over considerably greater distances, typically exceeding tens of meters. RFID is an additional Radio Frequency (RF) method used for estimation-based indoor positioning. RFID tags help with indoor locating by reflecting the RF signal after adding extra information.

RFID systems generally have eight power levels for RFID readers. The reader calculates the relevant power level for each tag based on the received signal strength, and repeats this method for all tag power levels. This method generates RSS data, which enables for the precise calculation of an item's position. Accuracy normally ranges between (15-20) cm. Given that infrared (IR) and RF-based localization systems may lack granularity or be expensive, there is need for more economical and finely granulated locating systems, such as acoustical systems, particularly those that use ultrasonic technology.

RFID is a highly robust indoor localization technique that can be used in various applications. However, it is also known to be an expensive technology. The fingerprinting localization approach that relies on RSSI can be used in indoor localization applications utilizing RFID [44].

2.3.3 Wireless Fidelity (Wi-Fi)

Wi-Fi is a popular technology used for indoor localization, owing to the widespread availability of Wi-Fi systems. In public, private, and commercial settings, Wi-Fi is commonly used to provide internet access and networking capabilities to a variety of devices [45]. It operates between the 2.4GHz and 5GHz bandwidth values.

However, the power consumption of WLAN systems is relatively higher. Wi-Fi has become an excellent choice for indoor localization because Wi-Fi chipsets are present on almost all modern PCs, smartphones, and other portable devices [46]. Even without the need for new hardware devices to be connected to current Wi-Fi access points or infrastructure, modest localization solutions with acceptable localization precision and accuracy might be built. Additionally, fingerprinting-based solutions could employ the Wi-Fi infrastructure as reference sites for signal collection. The abovementioned RSSI, TOF that includes the TOA, RTT, and TDOA, AOA, or hybrid approaches can be used to create Wi-Fi-based localization services.

2.3.4 Bluetooth

Bluetooth is today a very popular wireless technology standard used to communicate data across short distances. It operates in the frequency range of (2.4-2.48) kHz. Because of its simplicity of implementation, Bluetooth has become a popular indoor location technique. Bluetooth-based locating systems rely largely on Received Signal Strength (RSS) methodologies, although other methods such as AoA and Time of Flight (ToF) may also be used with Bluetooth. One famous system, iBeacon, uses Bluetooth technology only for localization reasons [47]. With Bluetooth Low Energy (BLE), the latest version, providing an enhanced data rate of 24 Mbps and coverage range of (70-100) m, the technology is highly energy-efficient compared to previous versions [48]. Although positioning techniques like AoA and ToA can be employed with BLE, most BLE-based indoor positioning systems (IPS) rely on RSSI due to its simplicity. However, implementing an IPS based on Bluetooth technology requires more base stations compared to other technologies.

2.3.5 ZigBee

ZigBee also known the IEEE 802.15.4 standard is a well-liked wireless communication technology. It specifies the physical and MAC layers for cost-effective, low rate personal and local area networks and allows for seamless communication between devices made by different manufacturers [49].

ZigBee based wireless devices operate in the (0.868, 0.915, and 2.4) GHz frequency bands, and the greatest data rate is 250 kbps. In order to establish a secure connection between devices, it offers two different types of keys: multiple link keys and network keys. It is perfect for devices and applications that need little data usage, high security, and a long battery life [50].

In an indoor setting, a ZigBee's signal range is between 20 and 30 meters. Measurements of RSSI are frequently utilized for the determination of the range between two ZigBee-capable devices. ZigBee might be successfully and well employed for localization. However, it is not commonly employed in consumer devices, and ZigBee is not popular for localization applications.

2.4 Localization Techniques

Some localization algorithms are reviewed in this section. These localization algorithms use the arrival time, arrival angle and phase of the received signal, as well as predicted channel station information to location objects inside buildings. In the following subsections, we will discuss the most often employed range-based localization techniques.

2.4.1 Time-Of-Arrival (ToA)

ToA is a commonly used technique in IP systems. It involves measuring the time where it takes for a signal to mobile from a source point to a destination point in order to determine the distance between them. The advantage of using ToA in indoor positioning is that it can provide accurate distance measurements, which in turn can lead to precise localization. One disadvantage of ToA is that it is highly susceptible to multipath propagation, where signals bounce off walls and objects in the environment, leading to inaccuracies in distance measurements. The equation for calculating distance using ToA is:

$$d = \frac{c \cdot t}{2} \tag{2.1}$$

where the speed of light is $c \approx 3 \times 10^8$m/s, t is the time of flight

For small size regions, ToA is preferable over the RSSI approach. However, synchronization and processing time have an impact on ToA distance measurement [51]. The symmetric double-sided two-way may be used to decrease time synchronization inaccuracy. This approach computes the standard error by analyzing numerous back-and-forth signal propagation attempts between nodes.

2.4.2 Time-Difference-Of-Arrival (TDoA)

Time Difference of Arrival is a method used in indoor positioning systems to calculate the position of an object or person based on the time difference of arrival of signals from multiple reference points. This method relies on measuring the difference in arrival time between the signals received at different reference points. The advantages of TDoA in indoor positioning include its ability to provide accurate positioning even in environments with obstacles and its independence from clock synchronization between the target and the reference points. However, TDoA also has some disadvantages. One disadvantage is that a TDoA-based positioning system requires the recording and cross-correlation of signal waveforms from at least four base stations, which can increase the complexity and computational requirements of the system [52]. The equation for TDoA positioning is as Equation (2.2), where $\Delta t = t_{received} - t_{transmitted}$ is the difference of the time of arrival at two different reference points, and c is the speed of light.

$$d = c \cdot (t_{received} - t_{transmitted}) \tag{2.2}$$

For TOA techniques to work the transmitter and receiver must be time synchronized. Even a microsecond microsecond synchronization error can result in a 300m error in location estimation.

2.4.3 Angle-Of-Arrival (AoA)

One of the prominent methods used in indoor positioning is Angle-of-Arrival technology [53]. AoA technology utilizes the principle of measuring the angle at which a radio signal arrives at multiple receiving antennas to determine the position of the user or object. AoA positioning has several advantages. First, it provides higher accuracy compared to other indoor positioning methods such as received signal strength indication or time of arrival. Second, AoA is less susceptible to interference and signal distortions caused by multipath propagation in indoor environments. Third, AoA positioning can work well in both line-of-sight and non-line-of-sight scenarios, making it versatile for different indoor environments. Nevertheless, AoA positioning also has some disadvantages. One of the main challenges is the requirement for precise time synchronization between the transmitting and receiving antennas, which can be difficult to achieve in practice. Additionally, AoA positioning typically requires a complex hardware setup with multiple antennas, which can increase the cost and complexity of implementation. Overall, AoA technology is a promising method for indoor positioning due to its high accuracy, resilience to signal distortions, and suitability for various indoor environments.

In AOA localization, to locate the object, the system must measure the angles θ_1 and θ_2 shown in Figure 2.1 as well as the distance from the BS to the localization target [54]. Given also the coordinates of the base stations as (x_1, y_1) and (x_2, y_2), the location of the target is gıven by:

$$y = \frac{y_2 \cdot \tan(\theta_2) - x_2}{\tan(\theta_2) - \tan(\theta_1)}, \quad x = y \cdot \tan(\theta_1) \tag{2.3}$$

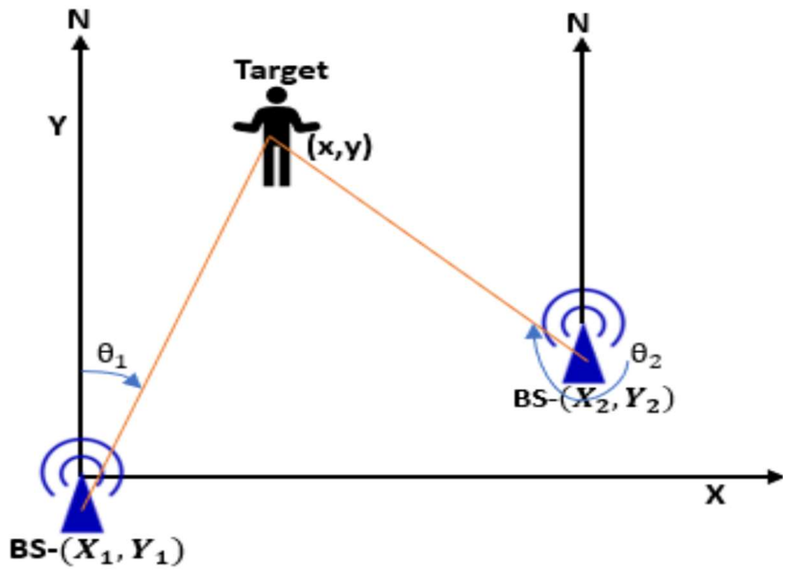

Figure 2.1. Smart Angulation-based localization measuring setup.

2.4.4 Received-Signal-Strength-Indicator (RSSI)

One of the commonly used methods in indoor positioning is the Received Signal Strength Indicator. The RSSI is a measurement of the power level of the received signal in wireless communication. So, it measures the strength of the received signal from wireless access points or routers to estimate the distance between the receiver and these reference points [55]. Although RSSI is simple and stable locating approach, it may generate incorrect distance measurements, particularly in small-scale situations, due to fading, interference, signal-shadowing, and scattering. As a result, ML techniques such as the ANN have been used to minimize RSSI fluctuations, and signal filters like Kalman filter (KF) may improve its functioning of the RSSI.

Using RSSI, as illustrated in Figure 2.2, to calculate the distance between transmitter and receiver requires a formula for math known as a channel model. The average RSSI should decrease with distance from the base station as described in Equation (2.4). Based on this equation, the distance between the transmitter and receiver can be estimated as shown in Equation (2.5):

$$RSSI_{io} = RSSI_o + 10 \cdot n \cdot \log_{10}\left(\frac{d_i}{d_0}\right) \tag{2.4}$$

$$d_i = d_o \cdot 10^{\left(\frac{RSSI_{io} - RSSI_i}{10n}\right)} \tag{2.5}$$

where n is a signal's propagating factor or exponent coefficient (which ranges from two in free space to four in interior conditions), d is the estimated distance in meters between the transmitter and the receiver, and $RSSI_o$ is the reference value of received RSSI at a distance of d_0 typically equal to 1 meter, and $RSSI_{io}$ is the current RSSI value at the receiver.

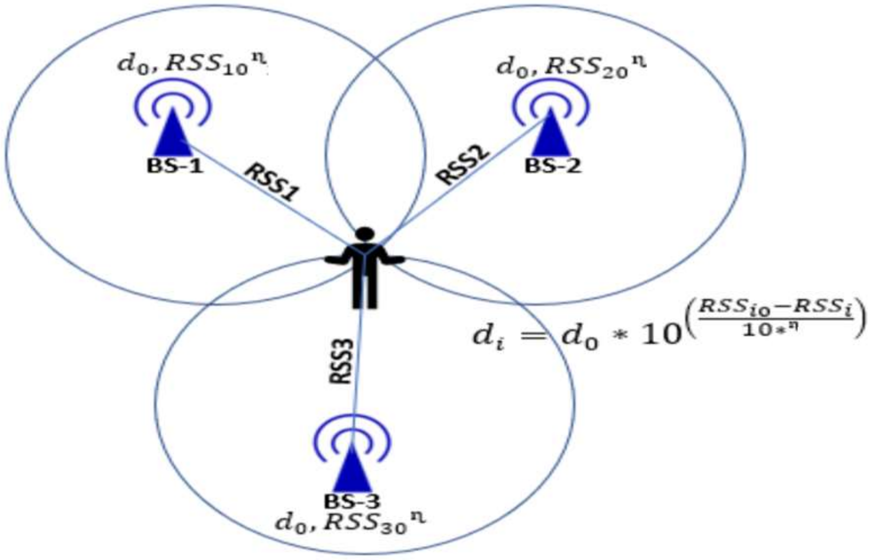

Figure 2.2. A distance calculations for a user with three stations system [56].

Advantages of using RSSI in indoor positioning include its low cost implementation and availability in most wireless communication devices, easy to connect and compatible to all wireless connection approaches. However, there are also disadvantages to using RSSI for indoor positioning. One disadvantage is that RSSI-based distance estimation provides poor accuracy, especially in non-ideal settings. Additionally, the RSSI values can be influenced by various factors such as multipath interference, signal fluctuations, and obstacles in the environment. As a result, it can employ additional support techniques that can optimize and enhance this approach like adopting the machine learning.

2.5 METHODS OF LOCALIZATION

In the field of positioning, there are various techniques used to determine the location of an object or individual. Trilateration is a technique that involves measuring the distances between an object and three known reference points to calculate its precise position in two or three-dimensional space. Sometimes, it uses both the distance and angle measurements from three or more reference points to determine the location of an object [57]. Multilateration, on the other hand, is a more general technique that involves using the distances from multiple known reference points to determine the position of an object, it often uses the time difference of arrival or time of flight of signals to determine the location [58][59]. This can be done by solving a system of equations based on the distances and the coordinates of the reference points. Fingerprinting is another positioning technique that relies on the unique characteristics of a specific environment. It relies on collecting and analyzing signal samples from the environment to create a unique fingerprint for each location, which can then be used to determine the location of an object or individual. For example, in indoor positioning systems, fingerprints of signal strengths from Wi-Fi access points or Bluetooth beacons are collected and stored in a database. When a user later enters the same environment, their current signal strengths can be compared to the database of fingerprints to estimate their position [37][60]. Global positioning systems are widely used for accurate positioning. They rely on a network of satellites that orbit the Earth and transmit signals. These signals are received by GPS receivers, which then use trilateration techniques to calculate the receiver's position based on the time it takes for the signals received [22]. Other localization methods, such as CoO, also known as Cell ID, involves determining the location based on the cell tower that is in closest proximity to the device. In summary, there are multiple localization methods available including trilateration, multilateration, fingerprinting, and CoO. These methods can be used individually or in combination to accurately determine the location of an object or individual.

2.6 MACHINE LEARNING FOR INDOOR LOCALIZATION

The indoor environment requires more accurate localization due to the presence of numerous obstacles and objects. Traditional outdoor localization methods, such as Global Positioning System, are not suitable for indoor environments due to signal attenuation and scattering caused by walls and other obstacles. Machine learning has emerged as a promising solution for indoor localization. ML algorithms have the ability to process large amounts of data and extract meaningful patterns, enabling accurate indoor localization. Various classifier types can be used in machine learning-based indoor localization, depending on the specific requirements and constraints of the application. Some common types of classifiers used in indoor localization include:

- K-nearest neighbour (KNN) classification algorithm: This algorithm classifies an object based on the majority vote of its nearest neighbours in the feature space.
- Neural networks (ANN): Neural networks are highly flexible and can learn complex patterns in data. They consist of interconnected nodes (neurons) that perform computations on input data and generate output predictions.
- Support vector machines (SVM): SVMs are powerful classifiers that separate data into various classes by finding an optimal hyperplane.

Leveraging machine learning (ML) techniques in indoor positioning systems (IPS) has become increasingly prominent for enhancing accuracy and robustness. ML algorithms analyze data patterns and make predictions, making them well-suited for addressing the complex and dynamic nature of indoor environments. Several types of machine learning approaches find application in IPS:

1. Supervised Learning:

I supervised learning, a data set is available with labels. The goal of the system is to learn how to reproduce the correct labels for new data into the system. The algorithm learns patterns and relationships between inputs, which are called features, and corresponding output labels. It can be employed for training models on datasets with known indoor locations, enabling the system to predict locations for new, unlabeled data.

2. Unsupervised Learning:

Unsupervised learning deals with datasets lacking predefined labels. This type of learning is more difficult, as the goals have to be stated clearly and there is more exploration that needs to be performed by the algorithm. The algorithm must aim to group the given data in some way or discover patterns in the data. For localization based on unsupervised learning, algorithms such as K-means as considered the significant clustering algorithm can be utilized to identify spatial patterns or group similar locations together in an indoor setting.

3. Reinforcement Learning:

In reinforcement learning, the learning is online and the algorithm learns to classify as it is processing data. After every decision, the algorithm receives feedback from the environment (which may be the user of the algorithm). Reinforcement learning can be employed to optimize the movement of mobile nodes within indoor spaces, learning optimal paths for accurate positioning.

4. Deep Learning:

Deep learning refers to a large neural network based algorithm, which is a type of machine learning algorithm which mimics the way neurons work. There are many types of deep learning systems, deep learning networks can use reinforcement learning, they can be used for both supervised and unsupervised learning. These networks are designed for advanced feature extraction and hierarchical learning. For IPS, deep learning models like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) find applications for extracting feature-rich representations and analyzing sequential data.

5. Semi-Supervised and Transfer Learning:

In semi-supervised learning some of the available training data has labels and some does not. Both data need to be processed by the system, depending on the application the semi-supervised system may or may not generate new data labels.

Transfer learning is when a learning algorithm trained for one task is used to achieve a different, usually related task. These approaches are beneficial when labeled data for indoor locations is limited, allowing models to generalize better from related tasks or partially labeled datasets.

6. Ensemble Learning:

Ensemble learning involves combining predictions from multiple models to enhance overall performance and robustness. Multiple ML models, each specialized in different aspects of indoor positioning, can be combined to provide more accurate and reliable results.

CHAPTER 3. MACHINE LEARNING ALGORITHMS AND PROPOSED WORK

3.1 Introduction

Traditional strategies for localization in indoor situations may be affected by many factors that restrict their operations and abilities to estimate an object position accurately and fast. Traditional localization techniques are ineffective in vast spaces like airports or retail malls. Furthermore, localization algorithms are not flexible to change or enlarge surroundings and disparate data sources. The benefit of machine learning (ML) algorithms is that they may learn relevant characteristics from applied input data. For instance, deep learning (DL) models can effectively study RSSI time series computations and use route knowledge to lessen RSSI oscillations [61]. One of the limitations of the fingerprinting localization approach is the existence of high-dimensional data. To address this, dimensionality reduction techniques like Principal Component Analysis (PCA) [62] can be employed to transform high-dimensional data into a less complex dimension, thereby reducing the computational and storage demands of fingerprinting-based localization.

Classification algorithms are typically utilized to extract key characteristics, for example, predefined positions (x, y, z) and signals powers (RSSI). Feature extraction using ML algorithms is especially useful for LOS and NLOS indoor localization. In practice, a substantial quantity of data characterizing the fingerprint map is acquired in the offline phase of fingerprint-based localization. As a result, estimating the online data using the fingerprint map data points takes time.

3.2 The Pre-processing

Collecting precise data from strategically positioned Access Points (APs) is essential in developing a robust Indoor Positioning System (IPS) based on Received Signal Strength Indicator (RSSI) powers. Deploying a network of APs allows for triangulation or fingerprinting and proximity estimation, forming the basis of accurate indoor localization. RSSI values, representing the signal strength between mobile devices (target) and APs, serve as fundamental proximity metrics in IPS.

Handling missing values in the dataset is a critical preprocessing step. Imputation techniques, such as mean complaint or more sophisticated machine learning-based methods, are employed to handle missing or gap values in the dataset caused by sporadic signal losses or environmental interferences. This ensures a comprehensive dataset for subsequent analysis.

The standardization of the dataset is paramount to normalize and removing biases in machine learning model training. Standardization like (z-score) method normalizes the data and transforms features to a common scale, mitigating issues related to varying measurement units and scales of different APs. This uniformity is crucial for the accurate functioning of machine learning algorithms.

Principal Component Analysis (PCA) is a robust method used for dimensionality reduction, especially in datasets with numerous attributes. By identifying and retaining essential information while discarding redundant features, PCA not only optimizes computational efficiency but also aids in visualizing the dataset's underlying structure.

Following preprocessing, the dataset undergoes a meticulous split into training and testing subsets. The training set, constituting 70% of the data, is utilized for model training, allowing algorithms to discern intricate patterns and relationships within the data. The remaining 30% serves as a testing subset, evaluating the model's generalization ability on unseen data.

The refined dataset, enriched through imputation, standardization, and potentially PCA, is now poised for integration with various machine learning classifiers. Decision Trees leverage hierarchical decision-making structures, while Support Vector Machines focus on finding optimal hyperplanes for classification. k-Nearest

Neighbors relies on proximity metrics, Logistic Regression models binary outcomes, Artificial Neural Networks simulate human neural networks, and Extreme Learning Machines employ efficient learning algorithms.

This amalgamation of scientific techniques ensures the IPS model's accuracy, reliability, and adaptability in diverse indoor environments, marking a significant stride in advanced indoor positioning technology.

3.3 Classification

Classification is the process of utilizing a training set of data that consists of observations (or attributes) with specified group identities to determine the set of classes (categories) to which a new observation belongs. The inputs for this process include the refined feature vector or set obtained after feature selection procedures and a classification dataset to be classified based on the aforementioned feature vector. The accuracy of detection algorithms is significantly influenced by the classification process [63]. Classification techniques can be broadly categorized into arithmetic, statistical, and intelligent approaches. Arithmetic methods involve numerical manipulations, such as KNN, LDA, LR, and SVM. Statistical methods rely on computing probability distributions and estimating parameters like mean and standard deviation to offer a more representative depiction of classes, as seen in DT. Intelligent methods, like Artificial Neural Networks (ANN), possess learning capabilities and employ artificial intelligence techniques in the classification process.

3.3.1 Decision Tree (DT)

The decision tree classifier is a machine learning algorithm that builds a predictive model in the form of a tree structure. Each node of the tree represents a decision or a splitting point based on a specific feature. These splitting points are determined through a process of evaluating different criteria, such as Gini impurity or information gain, to find the optimal way to divide the data. The decision tree classifier recursively splits the data based on these split points, creating branches and leaf nodes that represent different outcomes or classes [64][65].

Entropy
$$H(D) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{3.1}$$

Gini impurity
$$Gini(D) = 1 - \sum_{i=1}^{n} (p_i)^2 \tag{3.2}$$

where p_i is the probability of class i in dataset D.

The information gain of classifying by some property is:

Information gain
$$IG(D, A) = H(D) - H(D|\, a) \tag{3.3}$$

$$H(D|\, a) = \sum_{v=1}^{V} P_a(v)H(D_v) = \sum_{v=1}^{V} \frac{|D_v|}{|D|} H(D_v) \tag{3.4}$$

where A is a candidate attribute to split on, D_v is the subset of data for which attribute A is equal to v, and V is the set of all possible values of A.

The goal is to perform

3.3.2 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis is a supervised machine learning algorithm used for classification tasks. It is a dimensionality reduction technique that projects the data onto a lower-dimensional space in order to maximize the separation between classes. LDA assumes that the data follows a Gaussian distribution and that each class has its own mean vector and covariance matrix. Once computed, the LDA algorithm ranks the variables according to their importance score [66]. Then, it calculates a discriminant function for each class, which is a linear combination of the variables. The LDA classifier then uses these discriminant functions to classify new data points into one of the predefined classes. The LDA classifier is a linear classification method that aims to find a linear combination of features that best separates data points belonging to different classes. The basic form of the LDA equation is as follows:

$$y = W^T x \tag{3.5}$$

where y is the transformed feature vector in the lower-dimensional space. W is the transformation matrix. X is the original feature vector.

In the context of Linear Discriminant Analysis, where the goal is often to find a linear combination of features that best separates classes, the equation can be extended:

$$y = W^T x + b \tag{3.6}$$

where b represents a bias term.

3.3.3 Support Vector Machine (SVM)

The Support Vector Machine (SVM) stands out as a robust machine learning algorithm utilized for both classification and regression tasks. SVM works by identifying a hyperplane that effectively separates data points belonging to different classes in the feature space. This hyperplane is strategically determined to maximize the margin, which is the distance between the hyperplane and the nearest data point of each class. SVM can handle non-linear relationships by transforming the original feature space into a higher-dimensional space, allowing for more complex decision boundaries. The kernel is a function that performs a projection operation on that higher dimensional space. SVM is effective in scenarios where the data is not linearly separable, and it strives to create an optimal decision boundary with maximum margin [67][68].

Linear Kernel $$K(x, x_i) = x \cdot x_i \tag{3.7}$$

where x and x_i are input vectors.

Polynomial Kernel $$K(x, x_i) = (x \cdot x_i + c)^d \tag{3.8}$$

where c is a constant and d is the degree. The Gaussian kernel performs a projection onto an infinite dimensional space:

Gaussian (RBF) Kernel $$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{3.9}$$

where σ is the kernel width.

3.3.4 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a versatile machine learning algorithm used for classification and regression tasks. KNN classifies data points based on the majority class of their k nearest neighbors. The "k" in KNN represents the number of neighboring data points considered in the classification. KNN is a non-parametric method, meaning it does not make explicit assumptions about the underlying data distribution. It is particularly effective when dealing with locally varying patterns and is sensitive to the choice of distance metrics, such as Euclidean or Manhattan distance. KNN is straightforward to implement and well-suited for datasets with discernible clusters [69].

Weighted distance
$$d_i = \arg\max_c \left[\sum_{i=1}^{k} w_i I(y_i = c) \right] \tag{3.10}$$

where I is the indicator function and w_i is the weight of the i-th neighbor.

3.3.5 Logistic Regression (LR)

Logistic Regression (LR) is a linear model used for binary classification tasks. Despite its name, logistic regression is employed for classification, not regression. LR models the probability that a given input belongs to a particular class using the logistic function. The logistic function transforms the linear combination of input features into a range between 0 and 1, representing the probability. Logistic regression is widely used due to its simplicity, interpretability, and efficiency. It can be extended to handle multiclass classification through techniques like one-vs-rest or one-vs-one. Logistic regression is suitable for scenarios where the relationship between input features and the target variable is assumed to be linear [70].

Logistic Function (Sigmoid)
$$(Y = 1) = \frac{1}{1 + \exp\left(-(\sum_{i=1}^{n} w_i x_i + b)\right)} \tag{3.11}$$

Log-Likelihood
$$L(w) = \sum_{i=1}^{n} \left[y_i \log\big(P(Y = 1)\big) + (1 - y_i) \log\big(P(Y \neq 1)\big) \right] \tag{3.12}$$

3.3.6 Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) belongs to the category of machine learning models, drawing inspiration from the structure and functionality of the human brain. ANNs are composed of interconnected nodes arranged in layers, encompassing an input layer, one or more hidden layers, and an output layer. The connections between nodes have associated weights that are adjusted during the training process. ANNs can capture complex relationships in data and are capable of learning intricate patterns. Training an ANN involves feeding it with input data, adjusting weights based on the prediction errors, and iteratively refining the model. ANNs are known for their ability to handle non-linear relationships and are widely used in various applications, including image recognition and natural language processing [71][72].

Feedforward equation
$$a_j^{(l)} = g\left(\sum_{i=1}^{n} w_{ij}^{(l)} a_i^{(l-1)} + ab_j^{(l)}\right) \tag{3.13}$$

Backpropagation weight update
$$\Delta w_{ij} = -\eta \frac{\partial e}{\partial w_{ij}} \tag{3.14}$$

Where e is the error, η is the learning rate.

Stochastic gradient descent
$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \eta \frac{\partial e^{(t)}}{\partial w_{ij}^{(t)}} \tag{3.15}$$

3.3.7 Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) is a type of neural network machine learning algorithm known for its efficiency and simplicity. ELM is primarily used for supervised learning tasks, such as classification and regression. ELM differs from traditional neural networks in that it randomly assigns weights to input nodes and only adjusts the weights of the output layer during training. This randomness in weight initialization accelerates the training process, making ELM particularly suitable for large datasets. Despite its simplicity in structure as shown in Figure 3.1, ELM often exhibits competitive performance compared to more complex models. It is employed in scenarios where rapid training and prediction are essential [72][73]. Online Sequential Extreme Learning Machine (OS-ELM) [73] is an extension of the basic ELM that allows for sequential learning with individual samples.

The structure of ELM neural network includes three layers. The input layer receives input features from the dataset. This input data is multiplied by random gains before being passed on to the hidden layer, and remain fixed during training. The hidden layer neurons are responsible for learning and extracting features from the input data. In OS-ELM, neurons can be added to this layer sequentially, as data arrives and is used to train the network. The output layer produces the prediction.

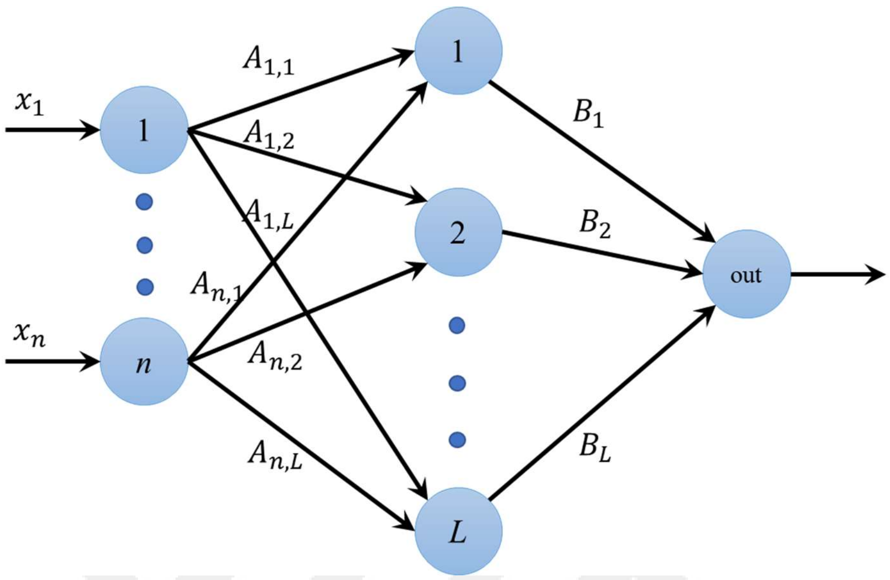

Figure 3.1. ELM basic Model.

OS-ELM updates its model parameters sequentially as new samples become available. This allows for online learning, making it suitable for scenarios where data arrives continuously. Similar to traditional ELM, OS-ELM exhibits fast learning due to the random initialization of weights in the hidden layer. These parameters do not need to be estimated, reducing the number of parameters the system needs to optimize.

Let us denote the key variables and parameters in OS-ELM. Let X be Input data matrix with dimensions $(n \times m)$, where n is the number of features, and m is the number of samples. Let H be Hidden layer output matrix with dimensions $(N \times m)$, where N is the number of hidden neurons. The weights from the input to the hidden layer form a random matrix with A with dimensions $(N \times n)$. B is the $(1 \times N)$ vector of weights from the hidden layer nodes to the output nodes. b_{in} is the vector of biases for the hidden layer with dimensions $(N \times 1)$. b_{out} is the vector of biases for the output layer.

The output of the hidden layer can be calculated for the i-th sample as:

$$H_i = g(AX_i + b_{in}) \tag{3.16}$$

where $g(\cdot)$ is the activation function (commonly a sigmoid or radial basis function).

The output of the network for the i-th sample is given by:

$$Y_i = BH_i + b_{out} \tag{3.17}$$

The weights A, B, and biases b_{in}, b_{out} are updated sequentially as new samples are introduced.

Training procedure for OS-ELM is as follows:

Initialization: Initialize weights A, B, and biases b_{in}, b_{out} with random values.

Sequential Learning: For each incoming sample X_i, update the hidden layer output H_i and update the model parameters using the sequential learning rule.

Prediction: After training, the model can be used to predict the output for new samples.

The Advantages of OS-ELM are that, firstly, it allows for continuous learning, making it suitable for scenarios with a continuous stream of data. It converges quickly due to its use of random weights between the first and middle layers. OS-ELM is also memory efficient, since it does not require the entire dataset to be stored in memory to perform batch functions.

3.4 Performance Evaluation Assessment

When assessing the effectiveness of classification algorithms, diverse measurement metrics come into play. A prevalent method involves employing a confusion matrix, also recognized as a contingency table. This table distinguishes four key categories: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP and TN denote accurately classified positive and negative instances, respectively. In contrast, FN signifies instances wrongly classified as negative but are, in fact, positive. Conversely, FP indicates instances wrongly classified as positive but are, in reality, negative. This research incorporates prevalent classification metrics, including precision, recall, F-Measure, accuracy, and specificity, offering insightful information on algorithm efficiency and facilitating comparative analysis.

While accuracy remains a prominent metric for evaluating classification algorithm performance, its utility diminishes in datasets with imbalanced classes. In such scenarios, metrics like precision and recall prove more fitting. Precision quantifies the number of relevant selected items from the total selected items, calculated by dividing true positives by the sum of true positives and false positives. In contrast, recall gauges the number of relevant selected items from the total relevant items, calculated by dividing true positives by the sum of true positives and false negatives. F-Measure amalgamates precision and recall into a singular score, capturing their harmonic mean. Specificity, as a final metric, gauges a test's ability to identify the absence of a condition. These metrics furnish a nuanced and accurate means of evaluating classification algorithms, particularly beneficial when confronted with imbalanced datasets.

These performance metrics offer a comprehensive evaluation, shedding light on the performance of classification algorithms:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + FN} \tag{3.18}$$

$$Precision = \frac{TP}{TP + FP} \tag{3.19}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.20}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.21}$$

$$F - measure = 2 \times \frac{precision \times recall}{precision + recall} \tag{3.22}$$

3.5 The proposed IP approach

Implementing an Indoor Positioning System (IPS) using WIFI that can be connected to a user easily using phone, laptop, or any connection device, with RSSI (in dbm) values collected from 5 APs to classify the target for 14 different places distributed non-uniformly in single floor place is our case under study. The single floor is moderate in size with area about 2600m^2, it is divided into 14 non-equal sections and rooms. 5 APs were employed sufficiently for collecting all required attributes to complete the fingerprinting position scan as shown in Figure 3.2.

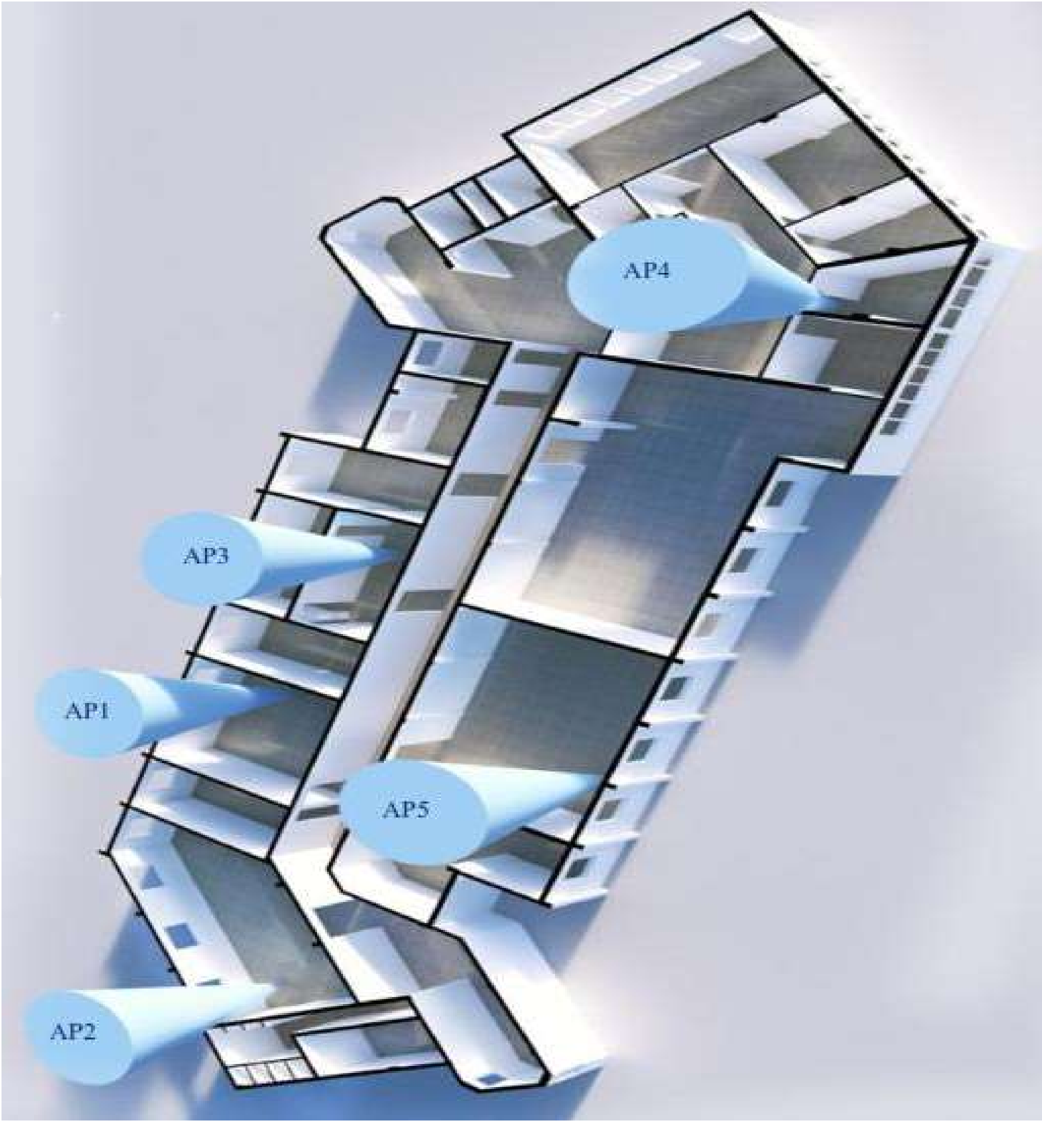

Figure 3.2. The diversity of the APs in the proposed floor map.

The collected data firstly is pre-processed and then applied to the well-known robust state of the arts classifier methods (DT, SVM, KNN, LDA, LR, and ANN) and to the Online Sequential Extreme Learning Machine (OS-ELM) classifier method involves several steps. The dataset is handled with scenarios, first without dimensionality reduction, second with PCA dimensionality reduction with neglecting 10% of the whole samples. Here's a general outline of the proposed process:

1. Data Collection: Collect RSSI values from the five access points (Aps) at known locations in the indoor environment. Associate each set of RSSI values with the corresponding ground truth location.

2. Data Preprocessing: Handle missing or noisy data in the RSSI measurements. Normalize or standardize the RSSI values to ensure consistency. If the gathered data is huge, we can employ the PCA to reduce the insignificant data to be able to minimize the required processing time.

3. Feature Extraction: Identify relevant features that can contribute to the accuracy of the positioning system. Extract additional information if available, such as angle of arrival or time-of-flight.

4. Dataset division: Split the dataset into training and testing sets with suitable ratios as 70% for the training data, and 30% for the testing.

5. Model Training: Implement five state of arts classifier methods, which are DT, SVM, KNN, LDA, ANN. Implement the Online Sequential Extreme Learning Machine (OS-ELM) algorithm by training using the preprocessed training data. Use the sequential nature of OS-ELM to continuously update the model as new data becomes available. Test the preprocessed testing data.

6. Model Evaluation: Evaluate the performance of the trained and test DT, SVM, KNN, LDA, ANN, and OS-ELM models on the testing set using appropriate metrics (e.g., Training accuracy, Testing accuracy, Training time, Prediction speed).

7. Fine-Tuning and Optimization: Iterate on the model and system performance, making adjustments as needed. Optimize hyperparameters, such as learning rates, number of hidden neurons, or regularization.

The comprehensive flow chart of the proposed method is as Figure 3.3, which includes two processes as:

1. **Offline Process:** The objective is to create a database by taking measurements at various locations within a predefined detection area sector. These measurements involve obtaining five RSS from five APs. The collected RSS power values from the APs are then linked to specific locations and classes within the maze.
2. **Online or Real time Process:** In this process, signal strength measurements are collected from multiple APs in the surrounding area. These measurements are then analyzed and grouped to ascertain the target's position within the designated location or maze. The strength of the signals from the APs serves as crucial

information in determining the target's location and class within the specified environment.

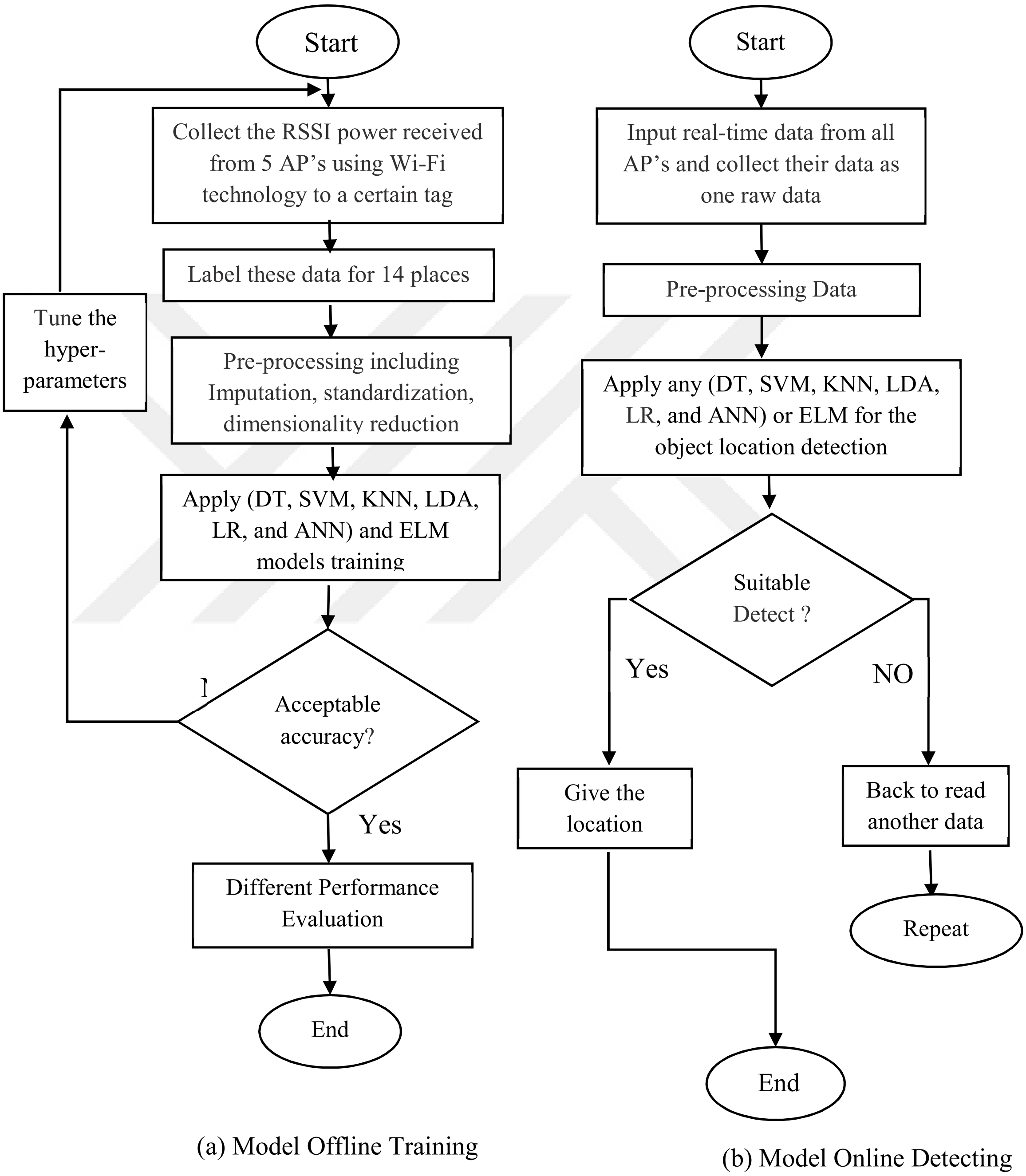

Figure 3.3. The proposed indoor positioning system flow chart.

CHAPTER 4. EXPERIMENTAL RESULTS

4.1 Introduction

In this chapter, the emphasis is placed on the practical implementation of IP detection through the utilization of various machine learning algorithms, as elucidated in the preceding chapters. The fundamental aspect involves the meticulous development of six distinct ML algorithms, namely Decision Tree (DT), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Logistic Regression (LR), and the Artificial Neural Network algorithm (ANN). Also, we develop extensively an Extreme Learning Machine called (OS-ELM).The ensuing results derived from both methodologies are thoroughly examined, fostering a comparative analysis between them and against pertinent findings from prior studies. To implement the proposed approach successfully, we need to utilize a combination of both software and hardware tools. For the software requirements, we were using the Matlab programming language (version 2022b) on Intel (R) Core i7-8550U CPU @1.80 GHz, 8 GB Random Access Memory (RAM), hard disk with a storage capacity of 500GB, and Intel(R) HD Graphics 520 MB.

4.2 The Applied Dataset

Exploiting the availability and facilities of the Wi-Fi technology, simplicity and low cost of 5 APs, with a user tag, the off-line process dataset that includes 700 samples by 6 features was collected. The first five features represents the RSSI powers in (dbm) for the 5 APs, while the last feature represents the target/class which is one of 14 different places. In brief, the dataset itself was divided into semi-equally 14 distinct categories that were labeled from 1 to 14. Table 4.1 shows some of these attributes.

Table 4.1 Some of the collected data attributes.

RSSI1	RSSI2	RSSI3	RSSI4	RSSI5	Target
-69	-57	-76	-100	-71	1
-67	-50	-78	-100	-77	1
-50	-58	-62	-100	-88	2
-51	-78	-68	-100	-79	2
-75	-63	-76	-100	-60	3
-71	-67	-83	-100	-64	3
-45	-78	-74	-100	-64	4
-36	-78	-68	-100	-69	4
-70	-81	-74	-75	-58	5
-72	-82	-66	-79	-55	5
-64	-82	-60	-90	-77	6
-65	-86	-59	-89	-80	6
-76	-100	-48	-100	-100	7
-77	-100	-45	-100	-84	7
-80	-100	-55	-86	-87	8
-82	-100	-52	-81	-82	8
-100	-100	-71	-69	-100	9
-100	-100	-74	-70	-100	9
-100	-100	-79	-51	-88	10
-100	-100	-82	-45	-85	10
-100	-100	-67	-57	-65	11
-100	-100	-73	-58	-67	11
-100	-100	-100	-56	-100	12
-100	-100	-98	-54	-100	12
-100	-100	-100	-60	-100	13
-100	-100	-100	-61	-100	13
-100	-62	-66	-63	-100	14
-100	-87	-66	-56	-100	14

The entire data was randomly divided into two main parts: training data with 70%, i.e, 490×6, while the remaining 30%, i.e, 210×6 was assigned to testing data. These data are evaluated and tested with 10 folded cross validation (10 CV).

4.3 Machine Learning Simulation Results

This section will showcase the simulation results obtained from each algorithm, with a subsequent in-depth discussion of the outcomes. The evaluation will specifically focus on time-related aspects and pertinent metrics employed to discern the efficacy of each algorithm. The entire analysis and presentation are conducted within the Matlab programming language. We had make the results in two scenarios as:

4.3.1 Classification without PCA Dimensionality Reduction

Here, all the data were considered without and dimensionality reduction, Figure 4.1 shows the whole systems used.

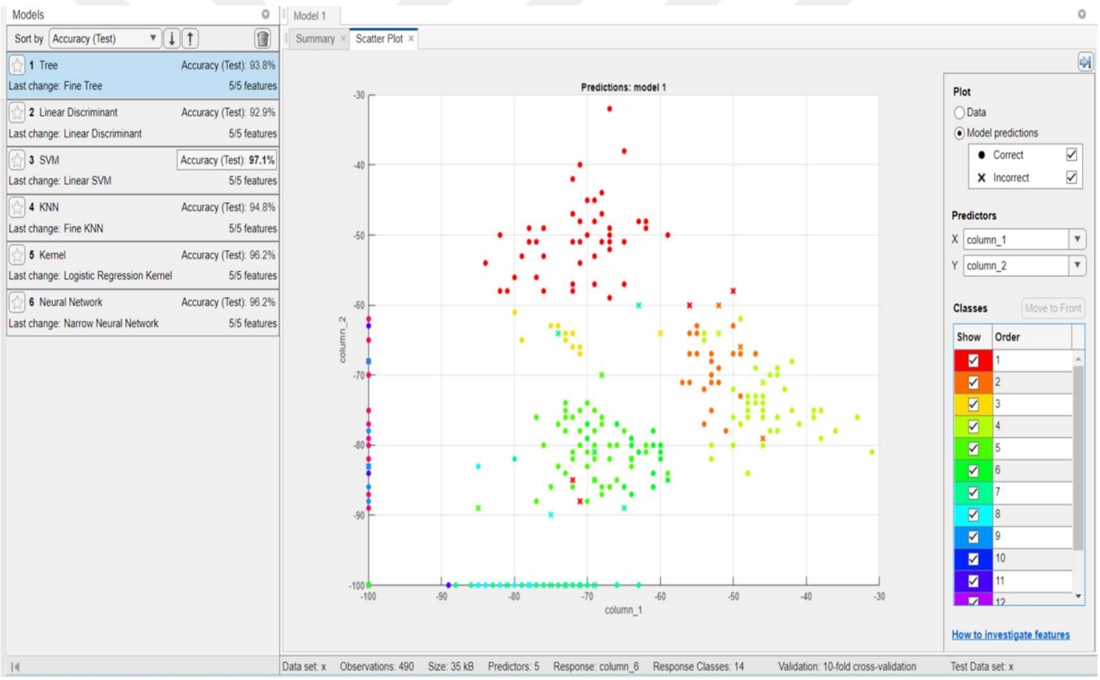

Figure 4.1. All machine learning used without PCA.

Figures 4.2 and 4.3 illustrate the DT classifier results, the training confusion matrix, and the testing confusion matrix.

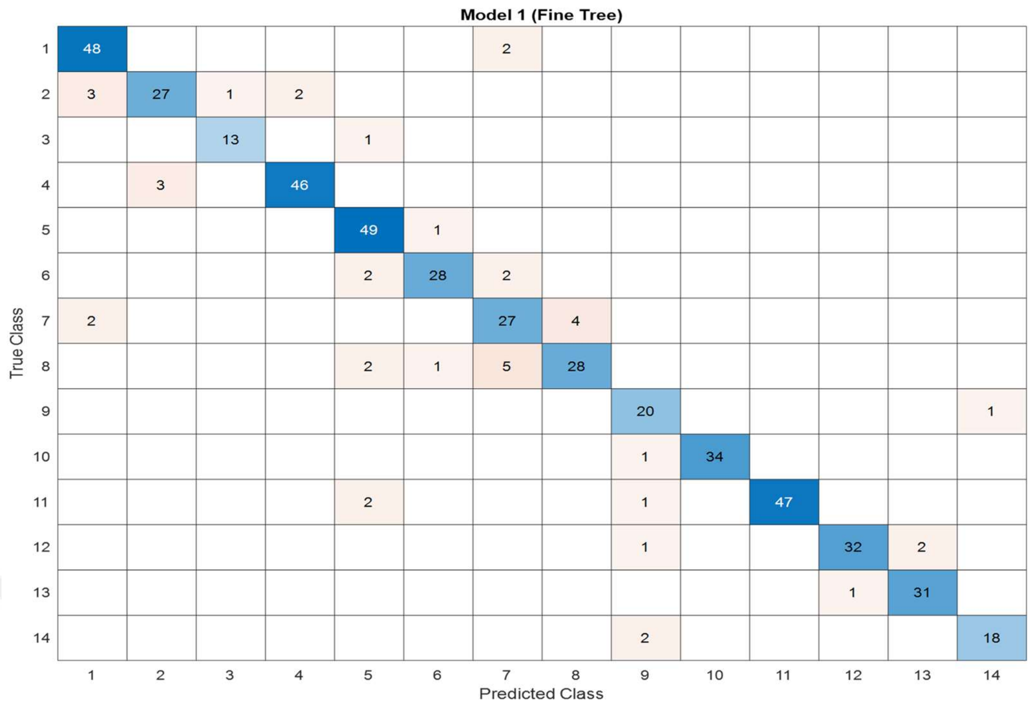

Figure 4.2. DT Confusion Matrix for training.

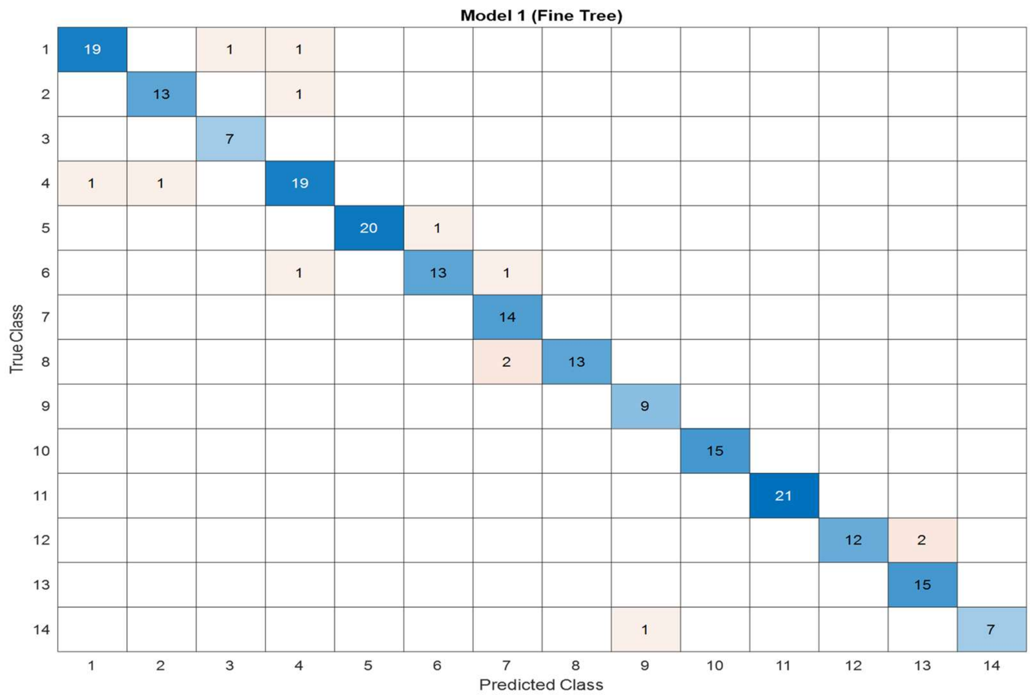

Figure 4.3. DT Confusion Matrix for testing.

Figures 4.4 and 4.5 depict the LDA classifier results, the training and the testing confusion matrices.

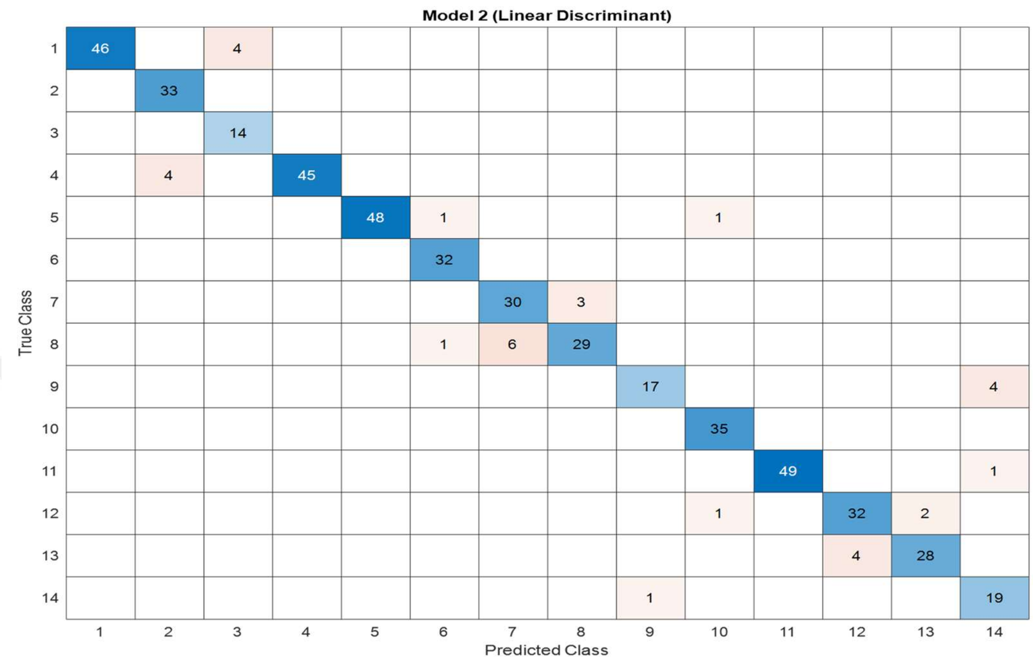

Figure 4.4. LDA Confusion Matrix for training.

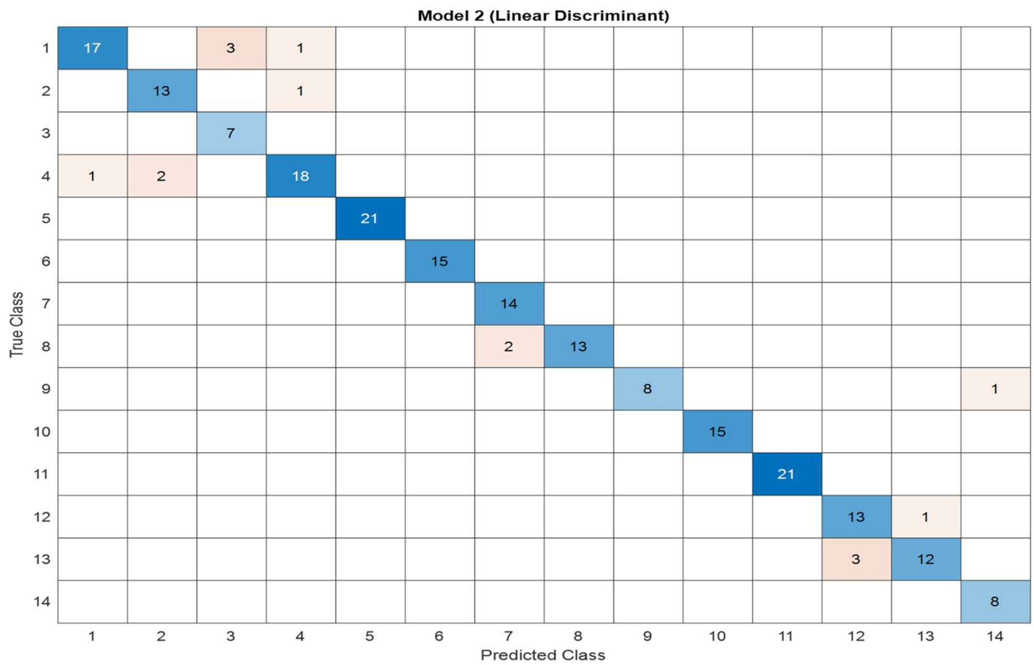

Figure 4.5. LDA Confusion Matrix for testing.

Figures 4.6 and 4.7 show the SVM classifier results, the training and the testing confusion matrices.

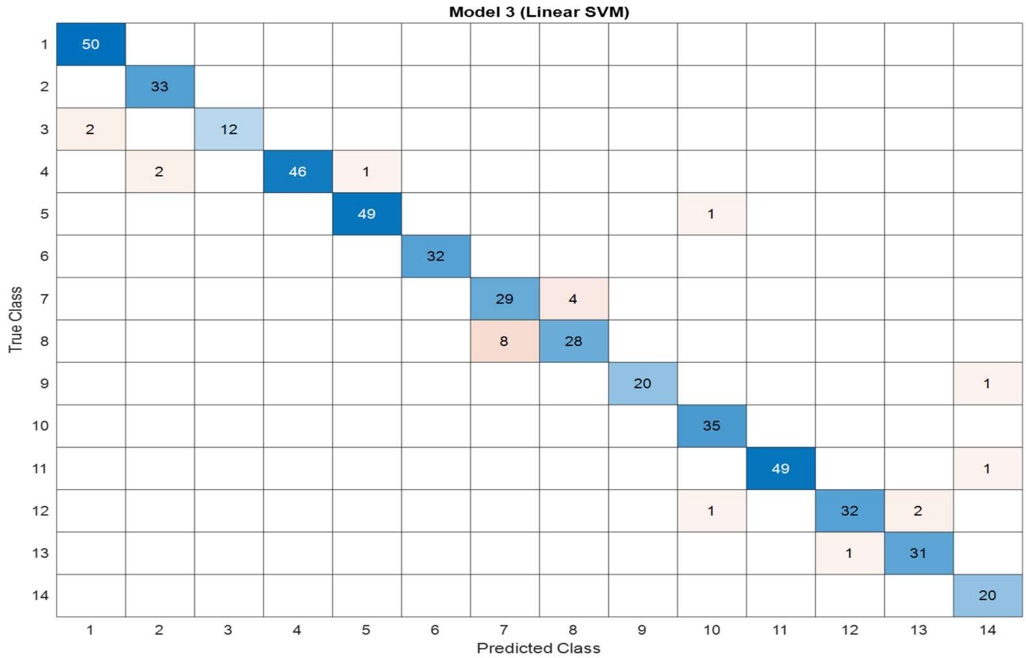

Figure 4.6. SVM Confusion Matrix for training.

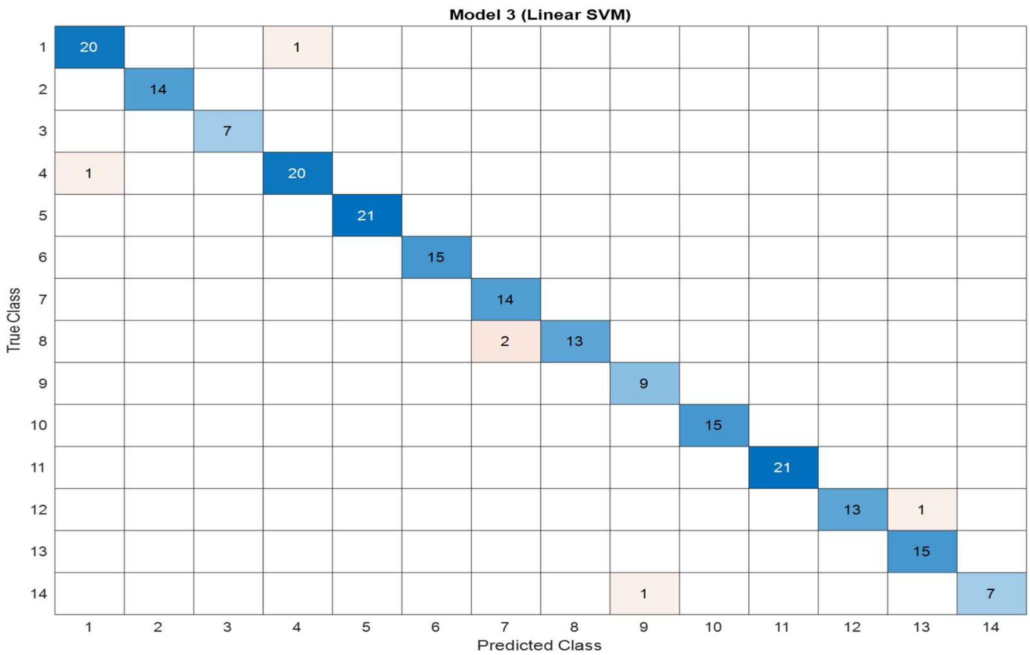

Figure 4.7. SVM Confusion Matrix for testing.

Figures 4.8 and 4.9 explain the LDA classifier results, the training and the testing confusion matrices.

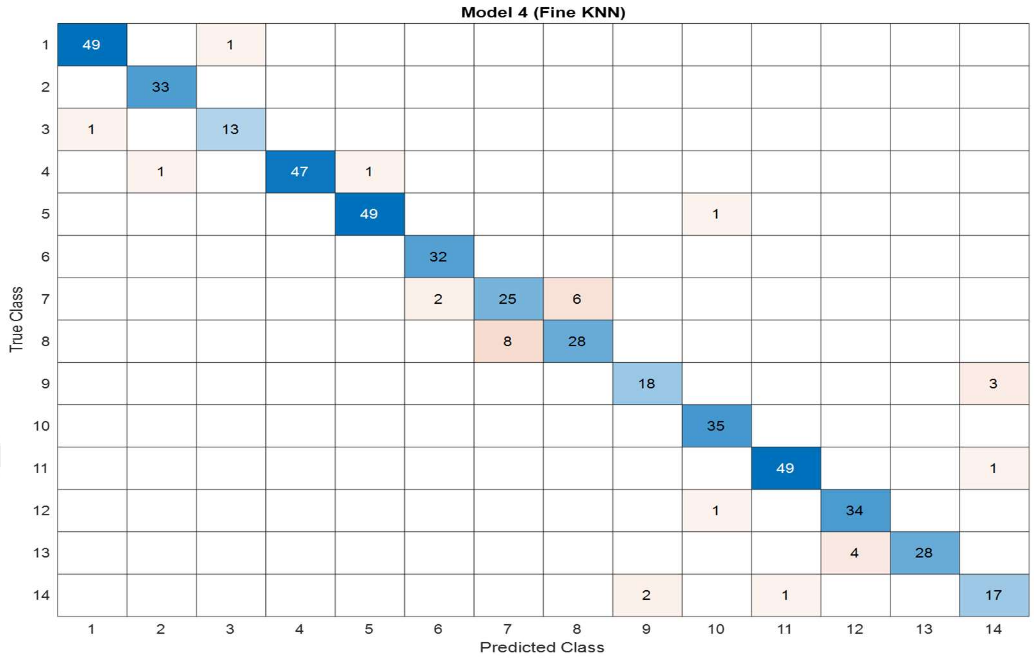

Figure 4.8. KNN Confusion Matrix for training.

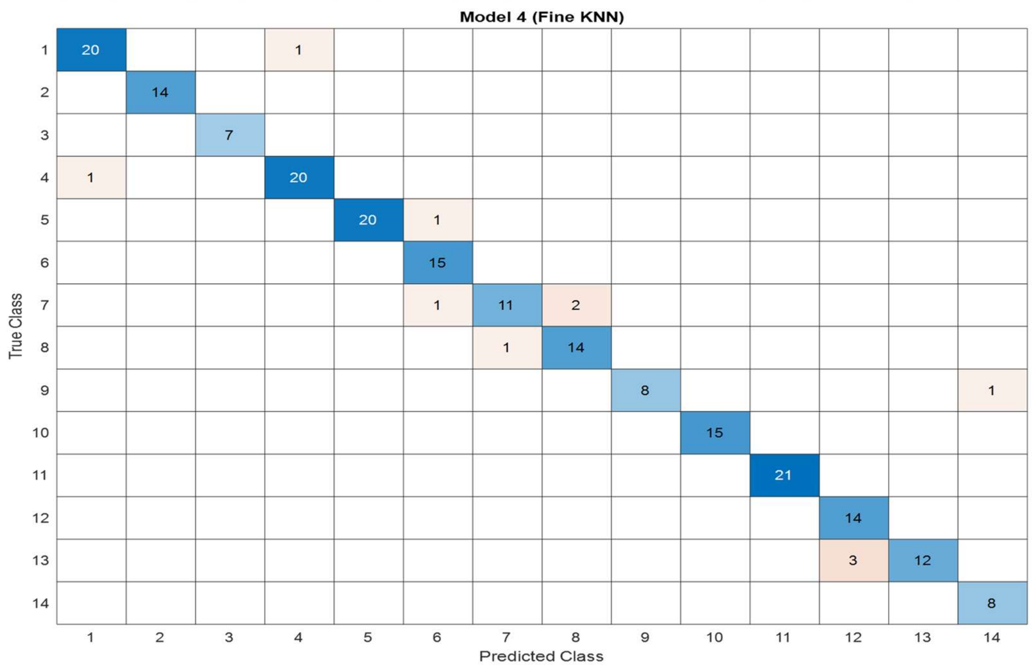

Figure 4.9. KNN Confusion Matrix for testing.

Figures 4.10 and 4.11 explain the LR classifier results, the training and the testing confusion matrices.

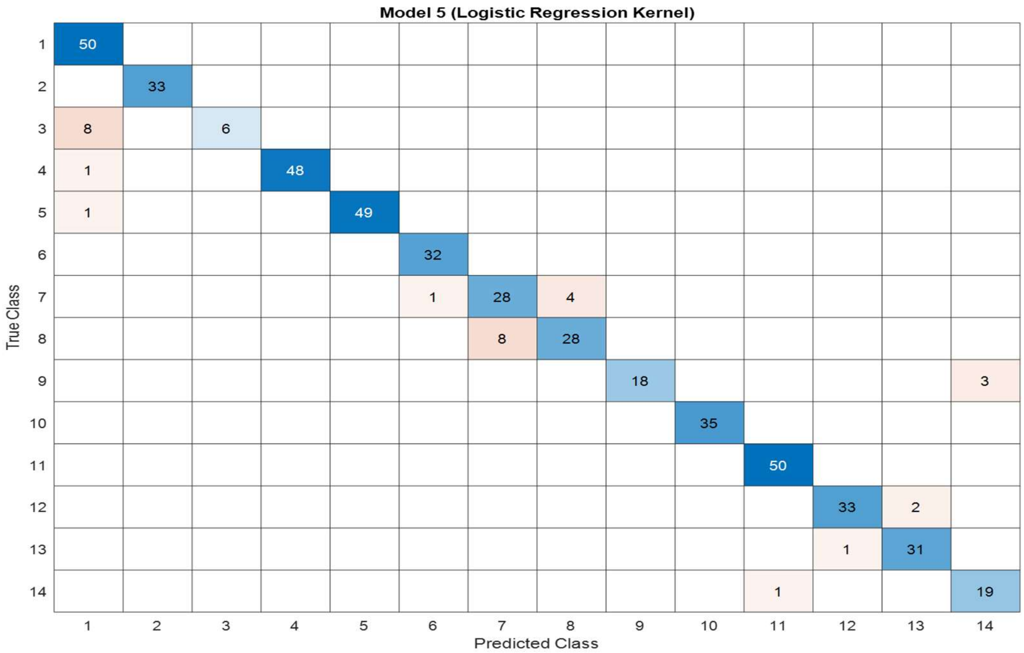

Figure 4.10. LR Confusion Matrix for training.

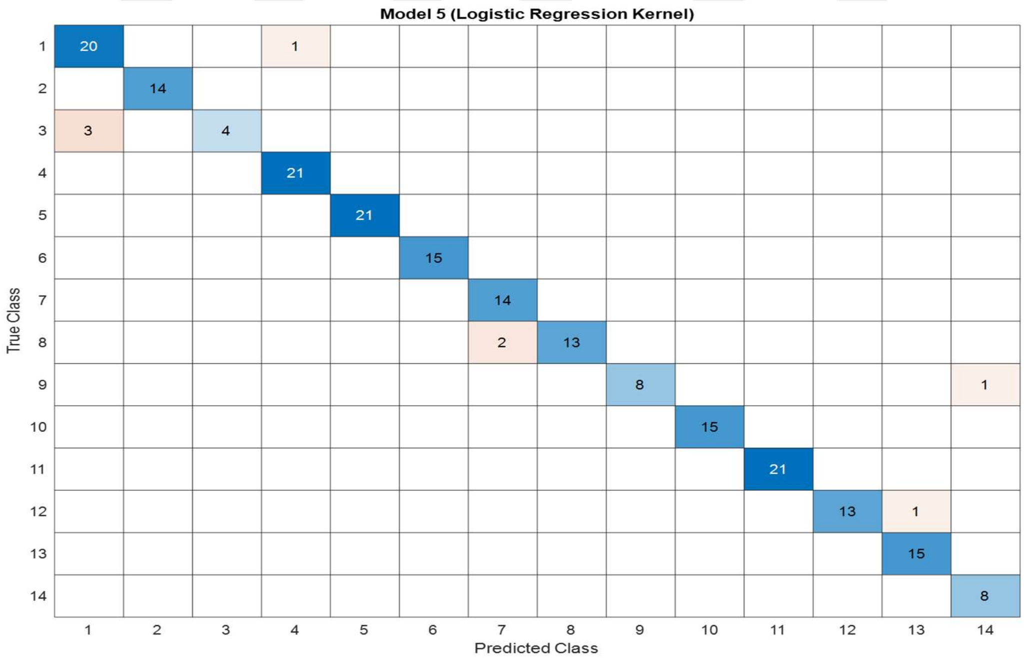

Figure 4.11. LR Confusion Matrix for testing.

Figures 4.12 and 4.13 illustrate the ANN classifier results, the training and the testing confusion matrices.

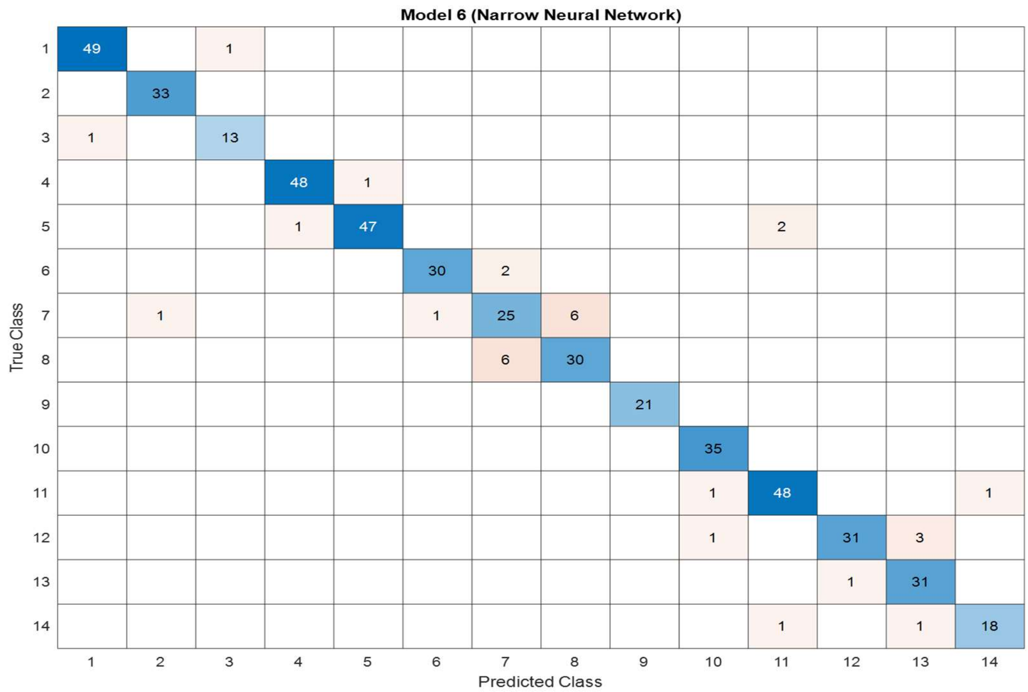

Figure 4.12. ANN Confusion Matrix for training.

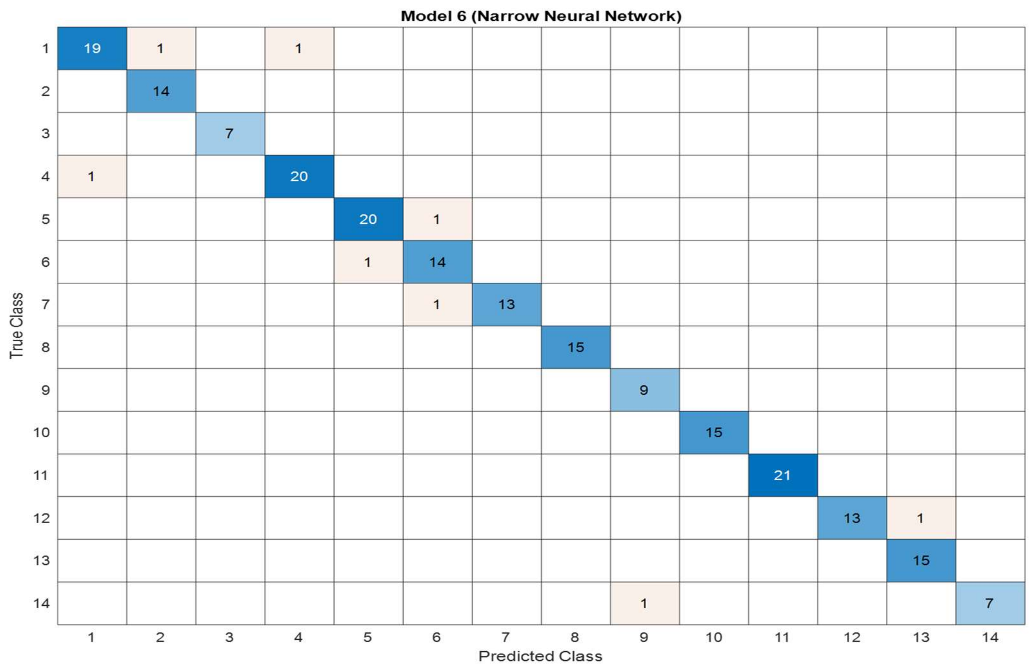

Figure 4.13. ANN Confusion Matrix for testing.

Now the ELM was executed with 10 cross validation and Nh=300 hidden layer nodes, as in table (4.2).

Table 4.2 ELM accuracies for 10 trials and their averages.

Train Accuracy (%)	Test Accuracy (%)
96.3576	89.6875
96.8543	95.1562
97.1854	94.0625
96.6887	99.5312
97.1854	95.1562
97.0199	89.6875
96.1921	96.2500
96.8543	86.4062
96.5232	94.0625
95.8609	92.9688
mean train accuracy = 96.6722%	mean test accuracy = 93.2969%
mean train time = 0.0516 sec	mean test time = 0.0078 sec

From the above table, it can be seen that the ELM has the best results including accuracy and time. Table (4.2), Figures 4.14 and 4.15 show all dataset splitting, classifiers information and resultant training and testing accuracies and times.

Table 4.3 All dataset, classifiers information and results.

Model No.	Model Type	Accuracy % (Validation)	Accuracy % (Test)	Prediction Speed (obs/sec)	Training Time (sec)
1	DT	91.429	90.810	10851.704	1.38
2	LDA	93.265	92.857	7770.476	1.27
3	SVM	95.102	93.143	1183.189	11.83
4	KNN	93.265	92.762	4425.230	2.14
5	LR	93.878	92.190	624.205	27.36
6	ANN	93.673	91.190	17914.464	11.95
7	ELM (Nh=300)	96.672	93.297	136500	0.0078

Table 4.5 shows the final results for classification without PCA dimensionality reduction. The total data is 490 samples, the data was separated into 210 test samples and 280 training samples. Since the data set is relatively small, the data is 10-fold cross-validated to ensure a good estimate of performance. The predictor is based on the first four predictor properties, with a total of 14 possible classes for the data to be classified into, corresponding to possible locations.

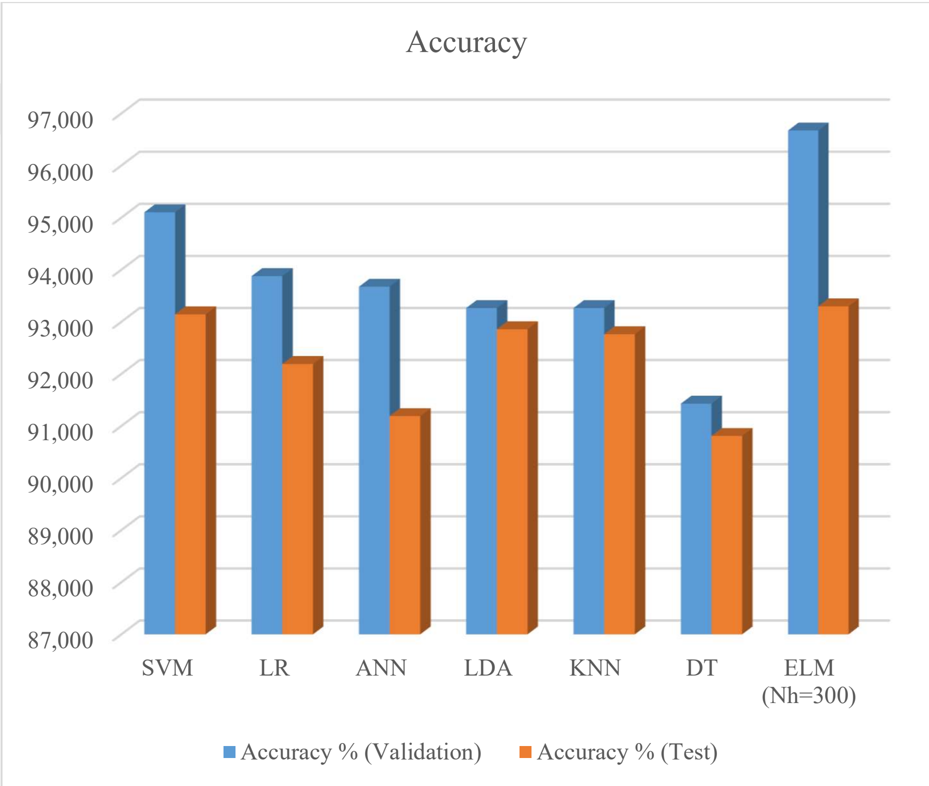

Figure 4.14. Training and testing accuracies.

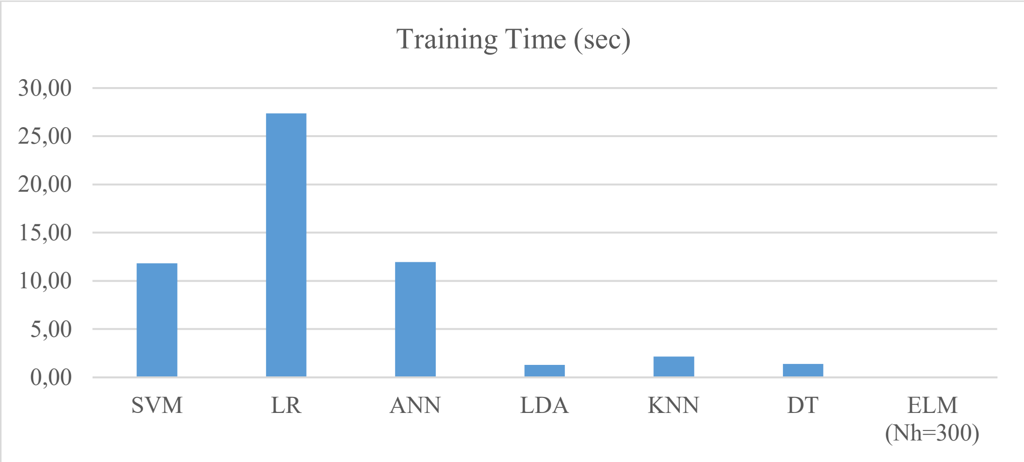

Figure 4.15. Training times.

In the classification analysis without PCA, the Extreme Learning Machine (ELM) with Nh=300 emerged as the top-performing classifier. It achieved the highest validation accuracy at 96.672% and maintained strong performance in the test dataset with an accuracy of 93.297%. Notably, ELM demonstrated remarkable prediction speed, registering 136500 observations per second, and an exceptionally low training time of 0.05 seconds.

On the other hand, the DT classifier, while still providing reasonable accuracy, exhibited comparatively lower performance. It secured a validation accuracy of 91.429% and a test accuracy of 90.810%. Additionally, the DT classifier showed a prediction speed of 10851.704 observations per second and a training time of 1.38 seconds.

In summary, ELM with Nh=300 stands out as the best-performing classifier, emphasizing high accuracy, rapid prediction speed, and swift training time. Conversely, the DT classifier, while functional, presents a slightly lower performance across these metrics.

4.3.2 Classification with PCA Dimensionality Reduction

Now, all the data were considered with PCA dimensionality reduction that reduced the data by 20%, so the resultant features size is 700×4 instead of 700×5, it can take into account the classifier results and explain how to be effected by this reduction.

Here, all the data were considered without and dimensionality reduction, Figure 4.16 shows the whole systems used.

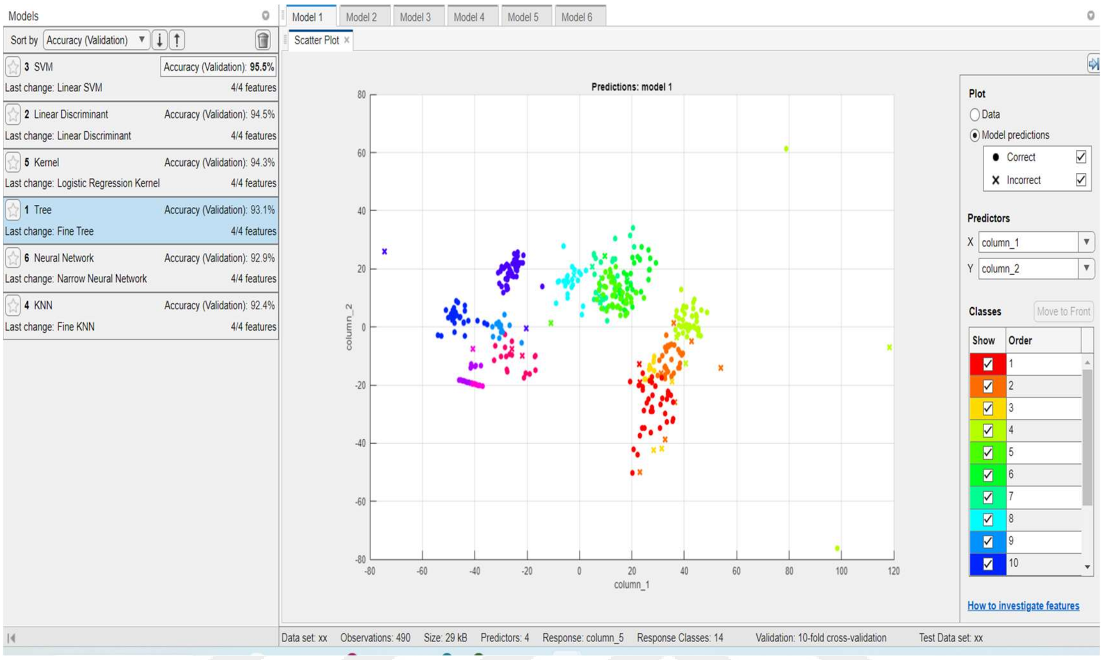

Figure 4.16. All machine learning used with PCA.

Figures 4.17 and 4.18 illustrate the DT classifier results, the training confusion matrix, and the testing confusion matrix.

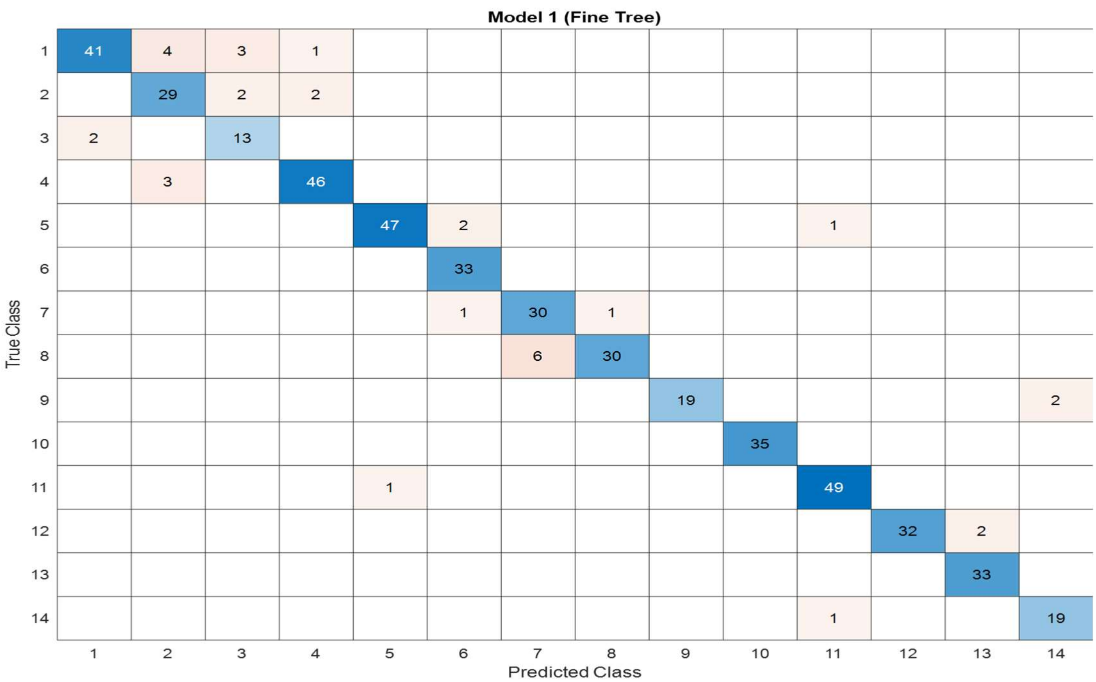

Figure 4.17. DT Confusion Matrix for training.

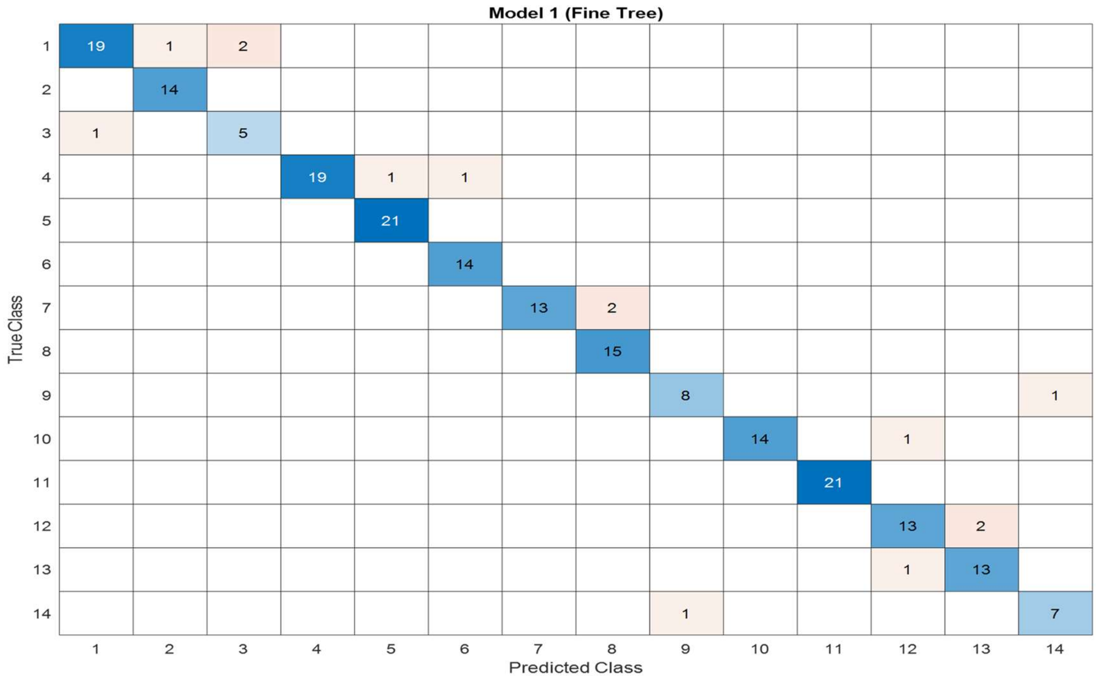

Figure 4.18. DT Confusion Matrix for testing.

Figures 4.19 and 4.20 depict the LDA classifier results, the training and the testing confusion matrices.

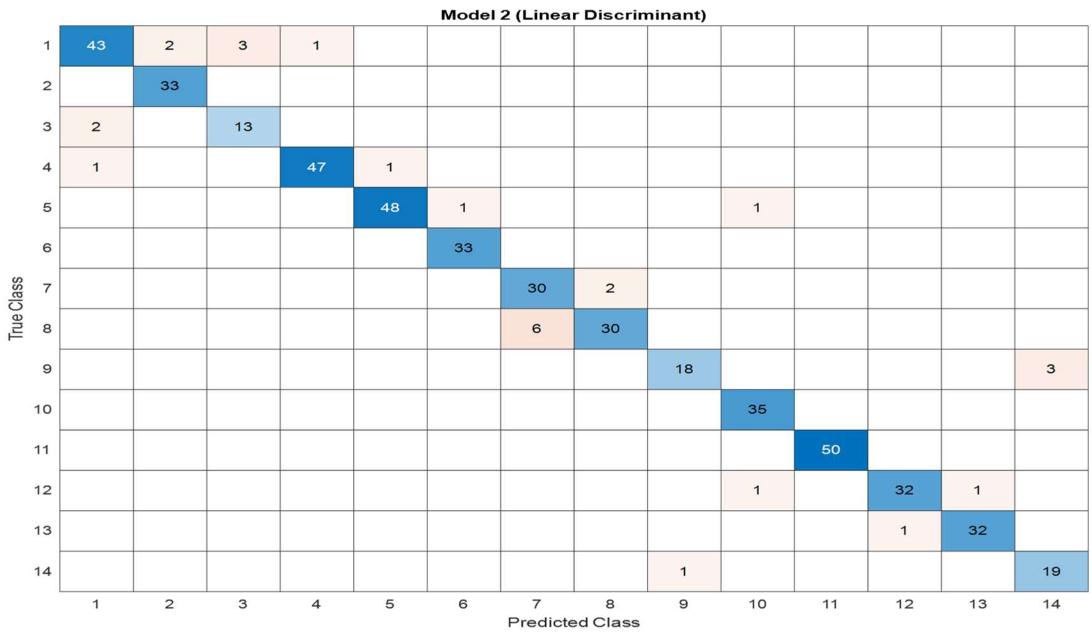

Figure 4.19. LDA Confusion Matrix for training.

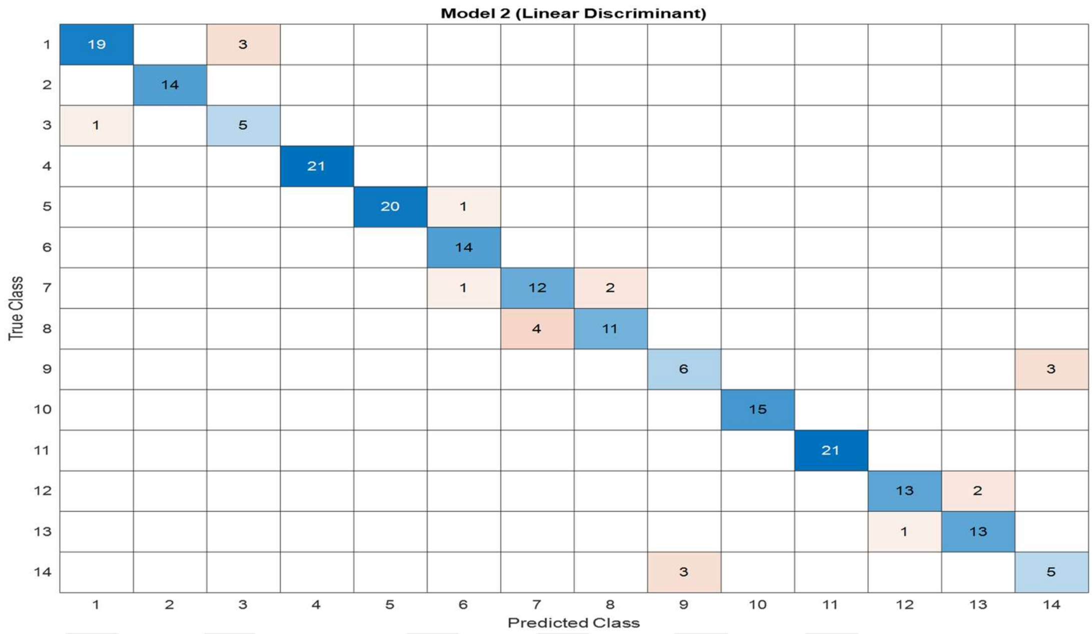

Figure 4.20. LDA Confusion Matrix for testing.

Figures 4.21 and 4.22 show the SVM classifier results, the training and the testing confusion matrices.

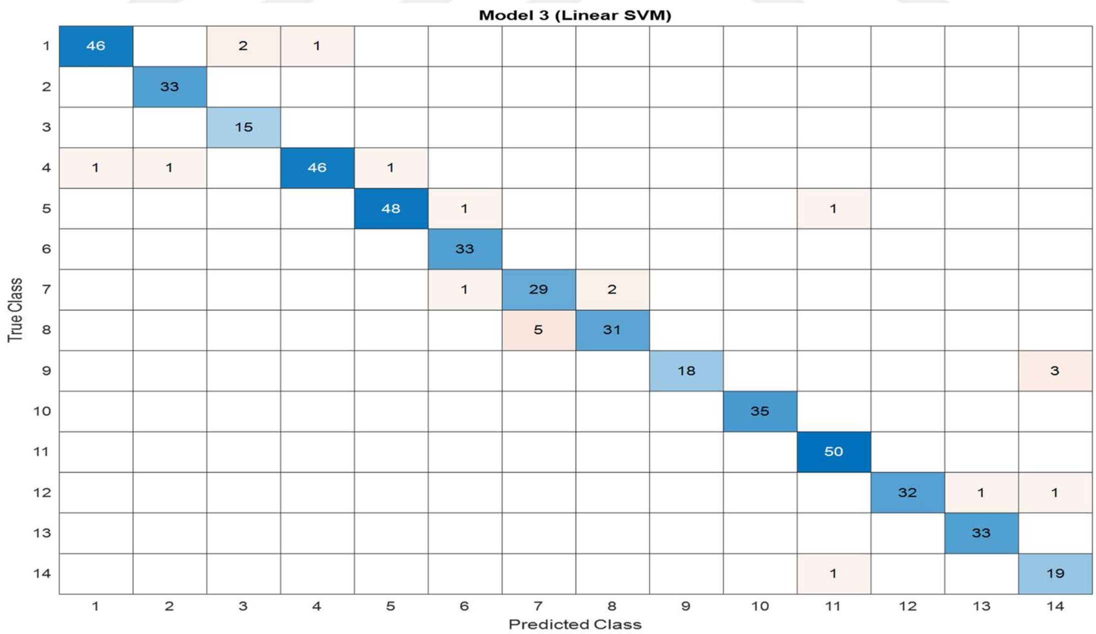

Figure 4.21. SVM Confusion Matrix for training.

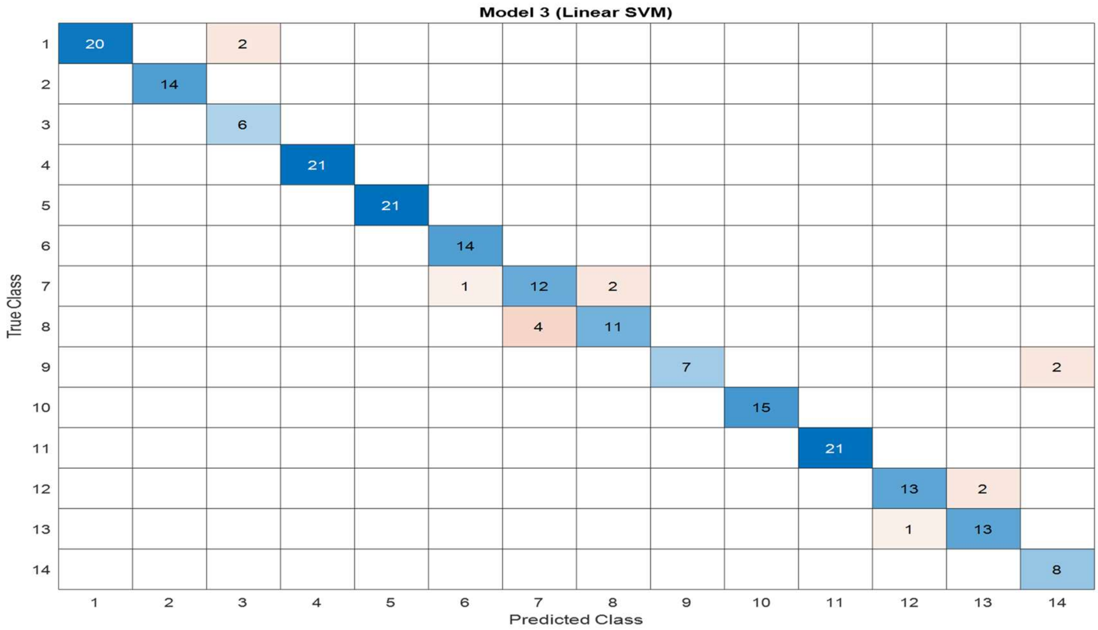

Figure 4.22. SVM Confusion Matrix for testing.

Figures 4.23 and 4.24 explain the KNN classifier results, the training and the testing confusion matrices.

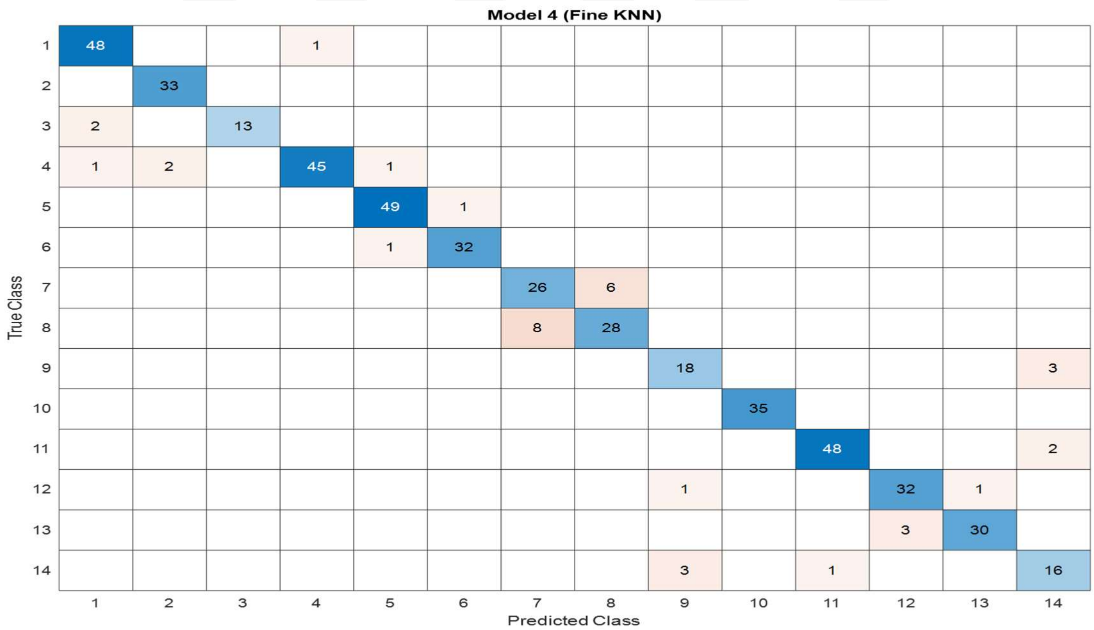

Figure 4.23. KNN Confusion Matrix for training.

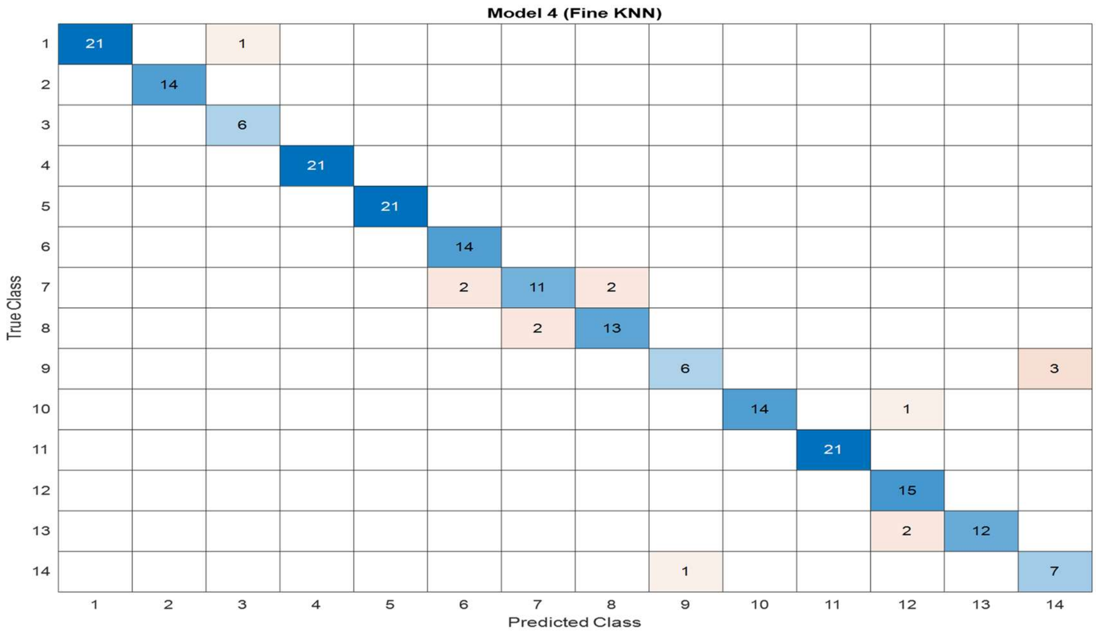

Figure 4.24. KNN Confusion Matrix for testing.

Figures 4.25 and 4.26 explain the LR classifier results, the training and the testing confusion matrices.

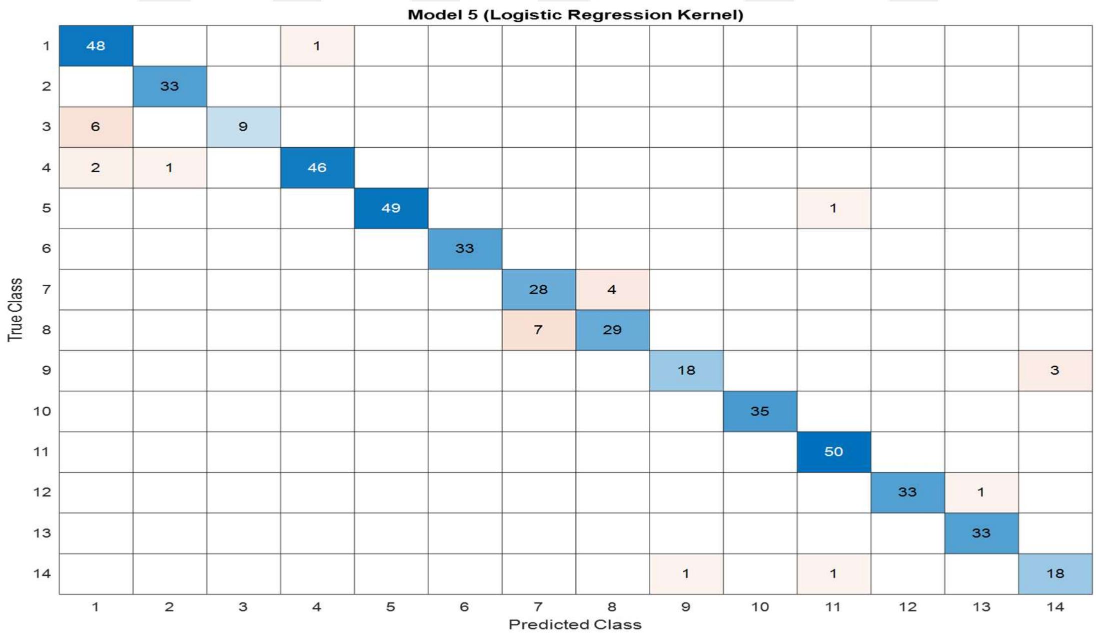

Figure 4.25. LR Confusion Matrix for training.

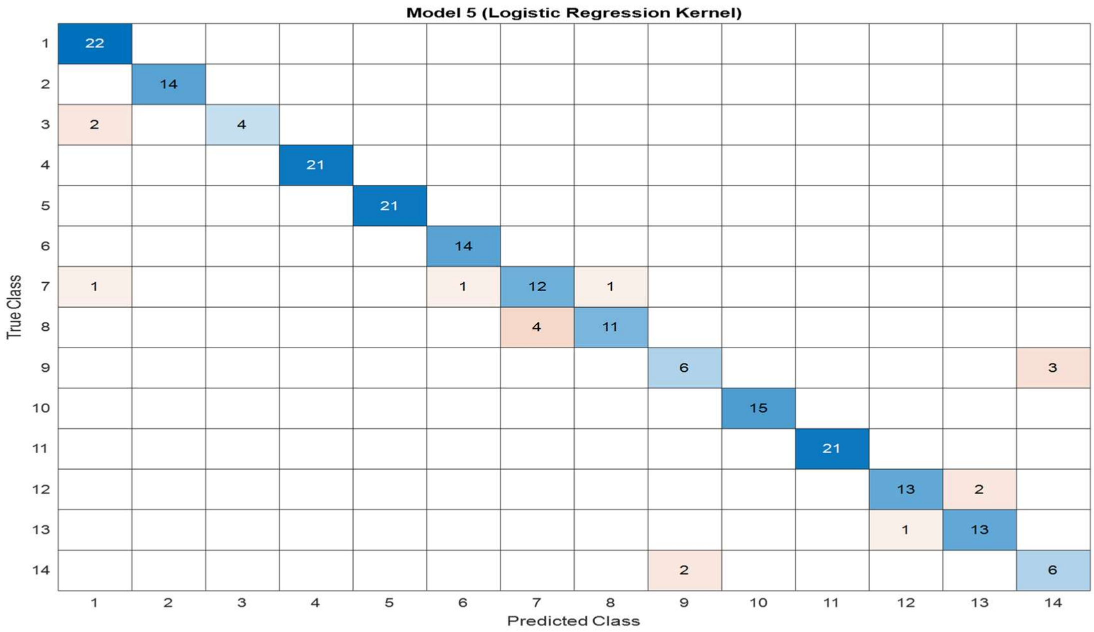

Figure 4.26. LR Confusion Matrix for testing.

Figures 4.27 and 4.28 illustrate the ANN classifier results, the training and the testing confusion matrices.

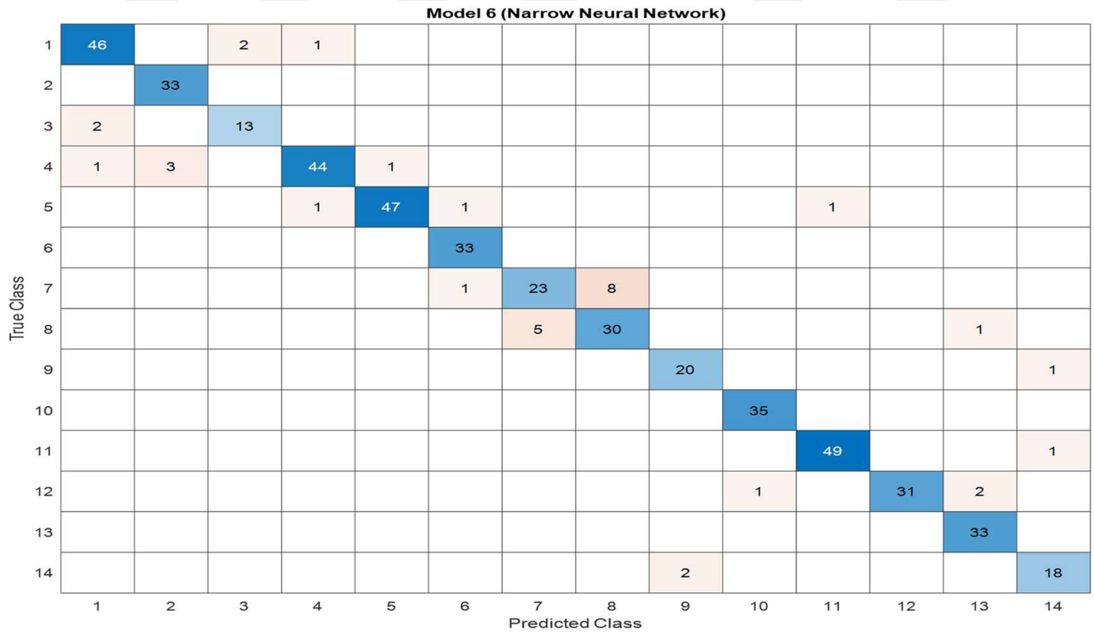

Figure 4.27. ANN Confusion Matrix for training.

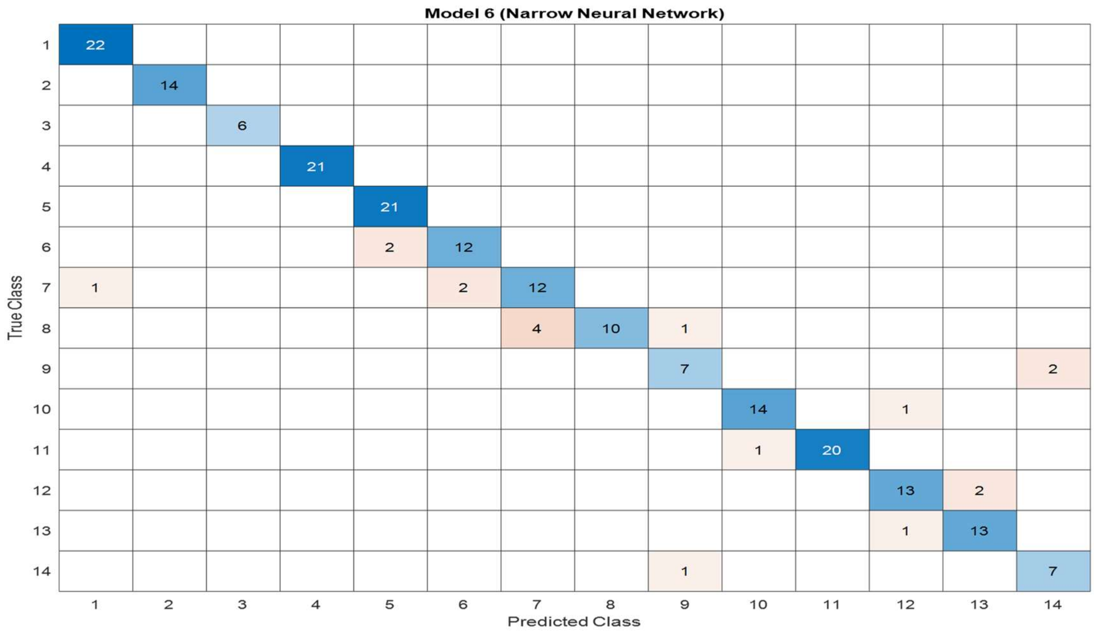

Figure 4.28. ANN Confusion Matrix for testing.

Now the ELM was executed with 10 cross validation and L=300 hidden layer nodes, as in table (4.4) below:

Table 4.4 ELM accuracies for 10 trials and their averages.

Train Accuracy (%)	Test Accuracy (%)
99.0066	91.8750
98.8411	92.9688
99.0066	91.8750
98.3444	96.2500
98.5099	97.3438
98.5099	95.1562
98.6755	98.4375
98.1788	99.5312
98.8411	94.0625
98.5099	95.1562
mean train accuracy= 98.6424%	mean Test Accuracy = 95.2656%
mean train time= 0.0434sec	mean test time= 0.0078 sec

From the above Figures, it can be seen that the ELM has the best results including accuracy and time. Table 4.5 and Figures 4.29 and 4.30 show all dataset splitting, classifiers information and resultant training and testing accuracies and times.

Table 4.5 shows the final results for classification with PCA dimensionality reduction. The total data is 490 samples, the data was separated into 210 test samples and 280 training samples. Since the data set is relatively small, the data is 10-fold cross-validated to ensure a good estimate of performance. The predictor is based on the first four predictor properties, with a total of 14 possible classes for the data to be classified into, corresponding to possible locations.

Table 4.5 All dataset, classifiers information and results.

Model No.	Model Type	Accuracy % (Validation)	Accuracy % (Test)	Prediction Speed (obs/sec)	Training Time (sec)
3	SVM	95.510	93.333	761.062	17.170
2	LDA	94.490	90.000	6256.863	1.825
5	LR	94.286	91.905	406.173	48.063
1	DT	93.061	93.333	6689.511	2.129
6	ANN	92.857	91.429	6505.585	24.054
4	KNN	92.449	93.333	2003.663	8.154
7	ELM (Nh=300)	98.6424	95.2656	157250	0.0078

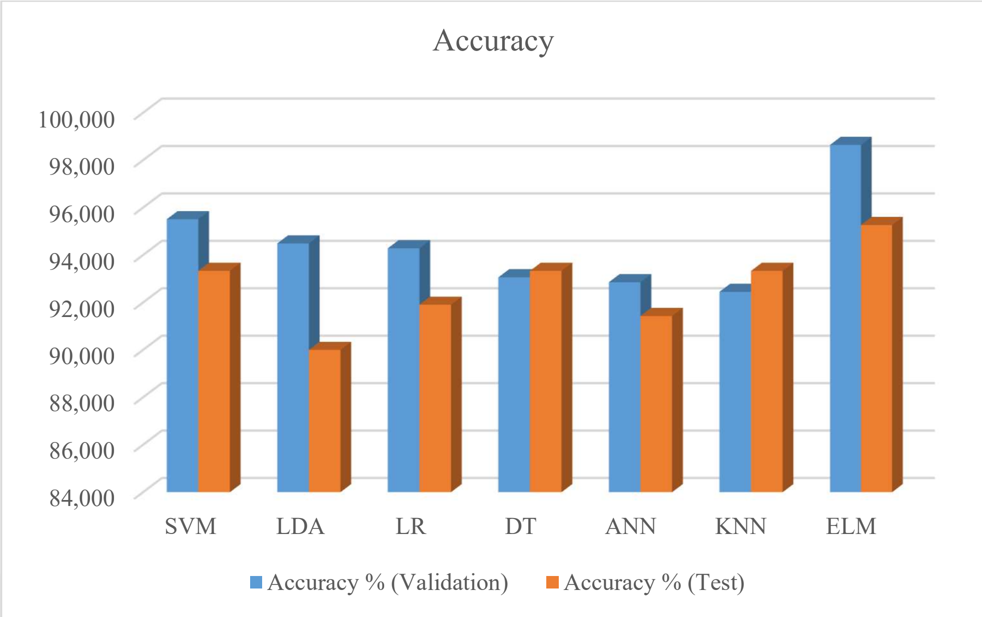

Figure 4.29. Training and testing accuracies.

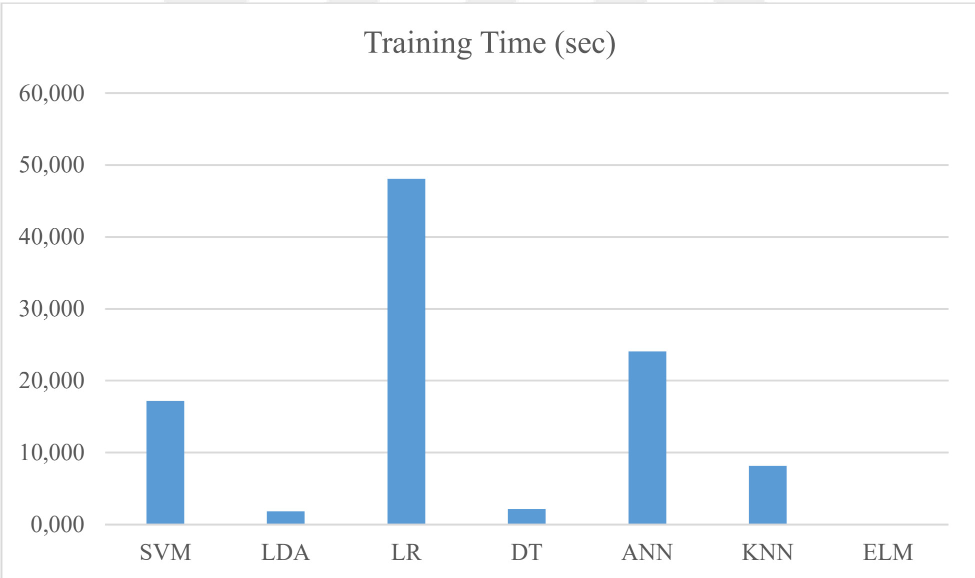

Figure 4.30. Training times.

In the classification analysis with PCA, the Extreme Learning Machine (ELM) outperformed other classifiers, particularly excelling in accuracy metrics. The ELM achieved an outstanding validation accuracy of 98.6424% and maintained exceptional accuracy in the test dataset at 95.2656%. Moreover, ELM showcased remarkable prediction speed, handling 157250 observations per second, and an astonishingly low training time of 0.0078 seconds.

Comparatively, the DT classifier again demonstrated solid performance but with slightly lower accuracy. It achieved a validation accuracy of 93.061% and a test accuracy of 93.333%. The DT classifier showed a prediction speed of 6689.511 observations per second and a training time of 2.129 seconds.

In summary, the ELM classifier continues to exhibit superior performance in the PCA-enhanced analysis, emphasizing exceptional accuracy, rapid prediction speed, and swift training time. The DT classifier, while reliable, presents a slightly lower accuracy and performance across the considered metrics.

CHAPTER 5. CONCLUSION AND FUTURE WORK

5.1 Introduction

The comprehensive exploration into indoor positioning (IP) using many distributed access points (Aps), Wi-Fi, user tag, machine learning (ML) algorithms has yielded insightful results, paving the way for advancements in accurate and efficient real-time localization. The seven distinct ML algorithms, namely Decision Tree (DT), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Logistic Regression (LR), Artificial Neural Network (ANN), and the innovative Extreme Learning Machine (OS-ELM), have undergone difficult evaluation. The focus has been on achieving optimal accuracy, prediction speed, and training time, ensuring applicability in diverse scenarios.

5.2 Conclusion

The implementation of IP detection using ML algorithms, notably the innovative OS-ELM, demonstrates remarkable effectiveness for indoor localization. Traditional strategies face challenges in large spaces, while ML algorithms offer adaptability and efficiency in diverse environments like the NLOS. The significance of dimensionality reduction techniques like PCA is underscored, addressing high-dimensional data in fingerprinting.

The numerical results showcase the exceptional performance of ML algorithms, with OS-ELM standing out in terms of accuracy, prediction speed, and training time. The synergy of different localization technologies, energy-efficient algorithms, and real-time implementations emerges as critical areas for future exploration. The conclusions are substantiated by real results, emphasizing the transformative potential of ML in indoor positioning.

- Accuracy: OS-ELM outperforms other classifiers, achieving the highest accuracy in both validation (98.64%) and test (95.27%) phases.

- Prediction Speed: LR exhibits the highest prediction speed (406.17 obs/sec), while OS-ELM, despite its exceptional accuracy, maintains a rapid prediction speed (157250 obs/sec).

- Training Time: OS-ELM excels with an astonishingly low training time (0.0078 sec), making it highly efficient for real-time applications.

5.3 Recommendations for Future Works

Future research should delve deeper into refining OS-ELM and exploring its adaptability to dynamic environments. Additionally, the implications of integrating multiple technologies and hybrid approaches for improved accuracy and robustness deserve thorough investigation.

1. Integration of Multiple Technologies: Investigate the collaboration of different localization technologies (Wi-Fi, UWB, etc.) for enhanced accuracy and robustness, substantiating the recommendations with specific research directions and potential outcomes.

2. Dynamic Learning Models: Explore dynamic ML models that can adapt to changing environments and user behaviors over time, providing insights into the feasibility and expected advantages.

3. Real-Time Implementation: Develop real-time implementations for ML algorithms, ensuring seamless integration into practical applications. Discuss potential challenges and considerations for real-world deployment.

4. Energy-Efficient Algorithms: Focus on developing energy-efficient ML algorithms for deployment in resource-constrained devices, delving into the significance of energy efficiency and potential implications.

5. Hybrid Approaches: Investigate hybrid approaches that combine fingerprinting with other localization techniques for improved performance, elucidating the rationale behind such combinations and expected benefits.

REFERENCES

[1] T. Arai, T. Yoshizawa, T. Aoki, K. Zempo, and Y. Okada, “Evaluation of Indoor Positioning System based on Attachable Infrared Beacons in Metal Shelf Environment,” in *2019 IEEE International Conference on Consumer Electronics (ICCE)*, IEEE, Jan. 2019, pp. 1–4. doi: 10.1109/ICCE.2019.8662007.

[2] S. El Abkari, J. El Mhamdi, A. Jilbab, and E. H. El Abkari, “ESP8266 Wireless Indoor Positioning System using Fingerprinting and Trilateration Techniques,” in *Lecture Notes in Networks and Systems*, vol. 211 LNNS, 2021, pp. 377–386. doi: 10.1007/978-3-030-73882-2_35.

[3] M. Saily, O. N. C. Yilmaz, D. S. Michalopoulos, E. Perez, R. Keating, and J. Schaepperle, “Positioning Technology Trends and Solutions Toward 6G,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, Institute of Electrical and Electronics Engineers Inc., Sep. 2021. doi: 10.1109/PIMRC50174.2021.9569341.

[4] I. Guvenc and C.-C. Chong, “A Survey on TOA Based Wireless Localization and NLOS Mitigation Techniques,” *IEEE Commun. Surv. Tutorials*, vol. 11, no. 3, pp. 107–124, 2009, doi: 10.1109/SURV.2009.090308.

[5] Y. Zheng, M. Sheng, J. Liu, and J. Li, “Exploiting AoA Estimation Accuracy for Indoor Localization: A Weighted AoA-Based Approach,” *IEEE Wirel. Commun. Lett.*, vol. 8, no. 1, pp. 65–68, Feb. 2019, doi: 10.1109/LWC.2018.2853745.

[6] Y. Xu, J. Zhou, and P. Zhang, “RSS-Based Source Localization When Path-Loss Model Parameters are Unknown,” *IEEE Commun. Lett.*, vol. 18, no. 6, pp.

1055–1058, Jun. 2014, doi: 10.1109/LCOMM.2014.2318031.

[7] J. Hu, "Wireless industrial indoor localization and its application," The Arctic University of Norway, 2017. [Online]. Available: https://munin.uit.no/handle/10037/11374

[8] R. Zhang, F. Hoflinger, and L. Reindl, "TDOA-Based Localization Using Interacting Multiple Model Estimator and Ultrasonic Transmitter/Receiver," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 8, pp. 2205–2214, Aug. 2013, doi: 10.1109/TIM.2013.2256713.

[9] M. Boltes, J. Adrian, and A.-K. Raytarowski, "A Hybrid Tracking System of Full-Body Motion Inside Crowds," *Sensors*, vol. 21, no. 6, p. 2108, Mar. 2021, doi: 10.3390/s21062108.

[10] A. Poulose and D. S. Han, "Hybrid Indoor Localization Using IMU Sensors and Smartphone Camera," *Sensors*, vol. 19, no. 23, p. 5084, Nov. 2019, doi: 10.3390/s19235084.

[11] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for WLAN location fingerprinting," in *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, 2004. MOBIQUITOUS 2004.*, IEEE, 2004, pp. 14–23. doi: 10.1109/MOBIQ.2004.1331706.

[12] Z. Yang, P. Zhang, and L. Chen, "RFID-enabled indoor positioning method for a real-time manufacturing execution system using OS-ELM," *Neurocomputing*, vol. 174, pp. 121–133, Jan. 2016, doi: 10.1016/j.neucom.2015.05.120.

[13] J. Zhang *et al.*, "Robust RFID Based 6-DoF Localization for Unmanned Aerial Vehicles," *IEEE Access*, vol. 7, pp. 77348–77361, 2019, doi: 10.1109/ACCESS.2019.2922211.

[14] B. Li, K. Zhao, and E. B. Sandoval, "A UWB-Based Indoor Positioning System Employing Neural Networks," *J. Geovisualization Spat. Anal.*, vol. 4, no. 2, p. 18, Dec. 2020, doi: 10.1007/s41651-020-00059-2.

[15] G. De Angelis, A. Moschitta, and P. Carbone, "Positioning Techniques in Indoor Environments Based on Stochastic Modeling of UWB Round-Trip-Time Measurements," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 8, pp. 2272–2281, Aug. 2016, doi: 10.1109/TITS.2016.2516822.

[16] S. Viswanathan and S. Srinivasan, “Improved path loss prediction model for short range indoor positioning using bluetooth low energy,” in *2015 IEEE SENSORS*, IEEE, Nov. 2015, pp. 1–4. doi: 10.1109/ICSENS.2015.7370397.

[17] C.-H. Cheng and S.-J. Syu, “Improving area positioning in ZigBee sensor networks using neural network algorithm,” *Microsyst. Technol.*, vol. 27, no. 4, pp. 1419–1428, Apr. 2021, doi: 10.1007/s00542-019-04309-2.

[18] M. A. Alanezi, H. R. E. H. Bouchekara, and M. S. Javaid, “Range-Based Localization of a Wireless Sensor Network for Internet of Things Using Received Signal Strength Indicator and the Most Valuable Player Algorithm,” *Technologies*, vol. 9, no. 2, p. 42, Jun. 2021, doi: 10.3390/technologies9020042.

[19] S. Djosic, I. Stojanovic, M. Jovanovic, T. Nikolic, and G. L. Djordjevic, “Fingerprinting-assisted UWB-based localization technique for complex indoor environments,” *Expert Syst. Appl.*, vol. 167, no. October 2020, p. 114188, Apr. 2021, doi: 10.1016/j.eswa.2020.114188.

[20] N. Chang, R. Rashidzadeh, and M. Ahmadi, “Robust indoor positioning using differential wi-fi access points,” *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1860–1867, Aug. 2010, doi: 10.1109/TCE.2010.5606338.

[21] A. Toyama, K. Mitsugi, K. Matsuo, E. Kulla, and L. Barolli, “Implementation of an Indoor Position Detecting System Using Mean BLE RSSI for Moving Omnidirectional Access Point Robot,” in *Lecture Notes in Networks and Systems*, vol. 278, 2021, pp. 225–234. doi: 10.1007/978-3-030-79725-6_22.

[22] V. Bui, N. T. Le, T. L. Vu, V. H. Nguyen, and Y. M. Jang, “GPS-Based Indoor/Outdoor Detection Scheme Using Machine Learning Techniques,” *Appl. Sci.*, vol. 10, no. 2, p. 500, Jan. 2020, doi: 10.3390/app10020500.

[23] Z. Zhang *et al.*, “An Enhanced Smartphone Indoor Positioning Scheme with Outlier Removal Using Machine Learning,” *Remote Sens.*, vol. 13, no. 6, p. 1106, Mar. 2021, doi: 10.3390/rs13061106.

[24] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, and K. Reddy, “Recurrent Neural Networks for Accurate RSSI Indoor Localization,” *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10639–10651, Dec. 2019, doi: 10.1109/JIOT.2019.2940368.

[25] H. Rizk and M. Youssef, “MonoDCell,” in *Proceedings of the 27th ACM*

SIGSPATIAL International Conference on Advances in Geographic Information Systems, New York, NY, USA: ACM, Nov. 2019, pp. 109–118. doi: 10.1145/3347146.3359065.

[26] Y. Hou, Y. Xue, C. Chen, and S. Xiao, "A RSS/AOA based indoor positioning system with a single LED lamp," in *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*, IEEE, Oct. 2015, pp. 1–4. doi: 10.1109/WCSP.2015.7341020.

[27] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 4, pp. 269–282, Sep. 2015, doi: 10.1145/2829988.2787487.

[28] H. Zou, X. Lu, H. Jiang, and L. Xie, "A Fast and Precise Indoor Localization Algorithm Based on an Online Sequential Extreme Learning Machine," *Sensors*, vol. 15, no. 1, pp. 1804–1824, Jan. 2015, doi: 10.3390/s150101804.

[29] N. T. Thuong, H. T. Phong, D.-T. Do, P. Van Hieu, and D. T. Loc, "Android application for WiFi based indoor position: System design and performance analysis," in *2016 International Conference on Information Networking (ICOIN)*, IEEE, Jan. 2016, pp. 416–419. doi: 10.1109/ICOIN.2016.7427147.

[30] Hongwei Ding, Zhengqi Zheng, and Y. Zhang, "AP weighted multiple matching nearest neighbors approach for fingerprint-based indoor localization," in *2016 Fourth International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services (UPINLBS)*, IEEE, Nov. 2016, pp. 218–222. doi: 10.1109/UPINLBS.2016.7809974.

[31] Y. E. Dari, S. S. Suyoto, and P. P. Pranowo, "CAPTURE: A Mobile Based Indoor Positioning System using Wireless Indoor Positioning System," *Int. J. Interact. Mob. Technol.*, vol. 12, no. 1, p. 61, Jan. 2018, doi: 10.3991/ijim.v12i1.7632.

[32] Y. Hou, X. Yang, and Q. Abbasi, "Efficient AoA-Based Wireless Indoor Localization for Hospital Outpatients Using Mobile Devices," *Sensors*, vol. 18, no. 11, p. 3698, Oct. 2018, doi: 10.3390/s18113698.

[33] O. A. Al-Ani, M. M. Abdulwahid, M. Mosleh, and R. A. Abd-Alhameed, "The Optimum Location for Access Point Deployment Based on RSS for Indoor Communication," *Int. J. Simul. Syst. Sci. Technol.*, no. c, Mar. 2019, doi:

10.5013/IJSSST.a.20.S1.02.

[34] M. Abbas, M. Elhamshary, H. Rizk, M. Torki, and M. Youssef, “WiDeep: WiFi-based Accurate and Robust Indoor Localization System using Deep Learning,” in *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*, IEEE, Mar. 2019, pp. 1–10. doi: 10.1109/PERCOM.2019.8767421.

[35] J. Xue, J. Liu, M. Sheng, Y. Shi, and J. Li, “A WiFi fingerprint based high-adaptability indoor localization via machine learning,” *China Commun.*, vol. 17, no. 7, pp. 247–259, Jul. 2020, doi: 10.23919/J.CC.2020.07.018.

[36] J. Yan, G. Qi, B. Kang, X. Wu, and H. Liu, “Extreme Learning Machine for Accurate Indoor Localization Using RSSI Fingerprints in Multifloor Environments,” *IEEE Internet Things J.*, vol. 8, no. 19, pp. 14623–14637, Oct. 2021, doi: 10.1109/JIOT.2021.3071152.

[37] J. Wang and J. Park, “An Enhanced Indoor Positioning Algorithm Based on Fingerprint Using Fine-Grained CSI and RSSI Measurements of IEEE 802.11n WLAN,” *Sensors*, vol. 21, no. 8, p. 2769, Apr. 2021, doi: 10.3390/s21082769.

[38] S. Djosic, I. Stojanovic, M. Jovanovic, and G. L. Djordjevic, “Multi-algorithm UWB-based localization method for mixed LOS/NLOS environments,” *Comput. Commun.*, vol. 181, no. April 2021, pp. 365–373, Jan. 2022, doi: 10.1016/j.comcom.2021.10.031.

[39] S. Shyam, S. J. Devaraj, K. Ezra, J. Delattre, and G. K. Lynus, “Design and implementation of UWB-based cyber-physical system for indoor localization in an industry environment,” in *Intelligent Edge Computing for Cyber Physical Applications*, Elsevier, 2023, pp. 167–185. doi: 10.1016/B978-0-323-99412-5.00010-1.

[40] K. A. Nguyen, Z. Luo, G. Li, and C. Watkins, “A review of smartphones-based indoor positioning: Challenges and applications,” *IET Cyber-systems Robot.*, vol. 3, no. 1, pp. 1–30, 2021, doi: 10.1049/csy2.12004.

[41] L. Cheng, Z. Wu, B. Lai, Q. Yang, A. Zhao, and Y. Wang, “Ultra Wideband Indoor Positioning System based on Artificial Intelligence Techniques,” in *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*, IEEE, Aug. 2020, pp. 438–444. doi:

10.1109/IRI49571.2020.00073.

[42] F. Qigao, S. Biwen, and W. Yaheng, “Tightly Coupled Model for Indoor Positioning based on UWB/INS,” *IJCSI Int. J. Comput. Sci. Issues*, vol. 12, no. 4, pp. 11–16, 2015.

[43] F. Zafari, A. Gkelias, and K. K. Leung, “A Survey of Indoor Localization Systems and Technologies,” *IEEE Commun. Surv. Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019, doi: 10.1109/COMST.2019.2911558.

[44] A. R. Jimenez Ruiz, F. Seco Granja, J. Carlos Prieto Honorato, and J. I. Guevara Rosas, “Pedestrian indoor navigation by aiding a foot-mounted IMU with RFID Signal Strength measurements,” in *2010 International Conference on Indoor Positioning and Indoor Navigation*, IEEE, Sep. 2010, pp. 1–7. doi: 10.1109/IPIN.2010.5646885.

[45] L. Tian, S. Santi, A. Seferagić, J. Lan, and J. Famaey, “Wi-Fi HaLow for the Internet of Things: An up-to-date survey on IEEE 802.11ah research,” *J. Netw. Comput. Appl.*, vol. 182, no. December 2020, p. 103036, May 2021, doi: 10.1016/j.jnca.2021.103036.

[46] S. Santi, T. De Koninck, G. Daneels, F. Lemic, and J. Famaey, “Location-Based Vertical Handovers in Wi-Fi Networks With IEEE 802.11ah,” *IEEE Access*, vol. 9, pp. 54389–54400, 2021, doi: 10.1109/ACCESS.2021.3071639.

[47] “Bluetooth.com.” Accessed: Dec. 01, 2024. [Online]. Available: https://www.bluetooth.com/

[48] V. Varshney, R. K. Goel, and M. A. Qadeer, “Indoor positioning system using Wi-Fi & Bluetooth Low Energy technology,” in *2016 Thirteenth International Conference on Wireless and Optical Communications Networks (WOCN)*, IEEE, Jul. 2016, pp. 1–6. doi: 10.1109/WOCN.2016.7759023.

[49] P. Baronti, P. Pillai, V. W. C. Chook, S. Chessa, A. Gotta, and Y. F. Hu, “Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards,” *Comput. Commun.*, vol. 30, no. 7, pp. 1655–1695, May 2007, doi: 10.1016/j.comcom.2006.12.020.

[50] Y. Li, “Positioning and Ranging Process of RSSI Algorithm Based on ZigBee Technology,” *J. Phys. Conf. Ser.*, vol. 1846, no. 1, p. 012082, Mar. 2021, doi: 10.1088/1742-6596/1846/1/012082.

[51] B. Song, S. Zhang, J. Long, and Q. Hu, “Fingerprinting Localization Method Based on TOA and Particle Filtering for Mines,” *Math. Probl. Eng.*, vol. 2017, no. 1, Jan. 2017, doi: 10.1155/2017/3215978.

[52] J. Schroeder, S. Galler, and K. Kyamakya, “A Low-Cost Experimental Ultra-Wideband Positioning System,” in *2005 IEEE International Conference on Ultra-Wideband*, IEEE, 2005, pp. 632–637. doi: 10.1109/ICU.2005.1570062.

[53] H. Zhang and Z. Zhang, “AOA-Based Three-Dimensional Positioning and Tracking Using the Factor Graph Technique,” *Symmetry (Basel).*, vol. 12, no. 9, p. 1400, Aug. 2020, doi: 10.3390/sym12091400.

[54] M. A. G. Al-Sadoon, R. Asif, Y. I. A. Al-Yasir, R. A. Abd-Alhameed, and P. S. Excell, “AOA Localization for Vehicle-Tracking Systems Using a Dual-Band Sensor Array,” *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 6330–6345, Aug. 2020, doi: 10.1109/TAP.2020.2981676.

[55] A. A. Sohan, M. Ali, F. Fairooz, A. I. Rahman, A. Chakrabarty, and M. R. Kabir, “Indoor Positioning Techniques using RSSI from Wireless Devices,” in *2019 22nd International Conference on Computer and Information Technology (ICCIT)*, IEEE, Dec. 2019, pp. 1–6. doi: 10.1109/ICCIT48885.2019.9038591.

[56] T. Li, S. Ai, S. Tateno, and Y. Hachiya, “Comparison of Multilateration Methods Using RSSI for Indoor Positioning System,” in *2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, IEEE, Sep. 2019, pp. 371–375. doi: 10.23919/SICE.2019.8859906.

[57] E. Teoman and T. Ovatman, “Trilateration in Indoor Positioning with an Uncertain Reference Point,” in *2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC)*, IEEE, May 2019, pp. 397–402. doi: 10.1109/ICNSC.2019.8743240.

[58] J. Yang, H. Lee, and K. Moessner, “Multilateration localization based on Singular Value Decomposition for 3D indoor positioning,” in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, Oct. 2016, pp. 1–8. doi: 10.1109/IPIN.2016.7743627.

[59] N. Miwa, S. Tagashira, H. Matsuda, T. Tsutsui, Y. Arakawa, and A. Fukuda, “A Multilateration-based Localization Scheme for Adhoc Wireless Positioning Networks Used in Information-oriented Construction,” in *2013 IEEE 27th*

International Conference on Advanced Information Networking and Applications (AINA), IEEE, Mar. 2013, pp. 690–695. doi: 10.1109/AINA.2013.52.

[60] K. Fong Peng Wye, S. Muhammad Mamduh Syed Zakaria, L. Munirah Kamarudin, A. Zakaria, N. Binti Ahmad, and K. Kamarudin, “RSS-based Fingerprinting Localization with Artificial Neural Network,” *J. Phys. Conf. Ser.*, vol. 1755, no. 1, p. 012033, Feb. 2021, doi: 10.1088/1742-6596/1755/1/012033.

[61] J. Bai, Y. Sun, W. Meng, and C. Li, “Wi-Fi Fingerprint-Based Indoor Mobile User Localization Using Deep Learning,” *Wirel. Commun. Mob. Comput.*, vol. 2021, pp. 1–12, Jan. 2021, doi: 10.1155/2021/6660990.

[62] L. Zhang, T. Tan, Y. Gong, and W. Yang, “Fingerprint Database Reconstruction Based on Robust PCA for Indoor Localization,” *Sensors*, vol. 19, no. 11, p. 2537, Jun. 2019, doi: 10.3390/s19112537.

[63] A. H. Salamah, M. Tamazin, M. A. Sharkas, and M. Khedr, “An enhanced WiFi indoor localization system based on machine learning,” in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, Oct. 2016, pp. 1–8. doi: 10.1109/IPIN.2016.7743586.

[64] Y. Nieto, V. Gacia-Diaz, C. Montenegro, C. C. Gonzalez, and R. Gonzalez Crespo, “Usage of Machine Learning for Strategic Decision Making at Higher Educational Institutions,” *IEEE Access*, vol. 7, pp. 75007–75017, 2019, doi: 10.1109/ACCESS.2019.2919343.

[65] N. A. Maung Maung, B. Y. Lwi, and S. Thida, “An Enhanced RSS Fingerprinting-based Wireless Indoor Positioning using Random Forest Classifier,” in *2020 International Conference on Advanced Information Technologies (ICAIT)*, IEEE, Nov. 2020, pp. 59–63. doi: 10.1109/ICAIT51105.2020.9261776.

[66] W. Zhang, X. Hua, K. Yu, W. Qiu, S. Zhang, and X. He, “A novel WiFi indoor positioning strategy based on weighted squared Euclidean distance and local principal gradient direction,” *Sens. Rev.*, vol. 39, no. 1, pp. 99–106, Jan. 2019, doi: 10.1108/SR-06-2017-0109.

[67] M. Petric, A. Neskovic, N. Neskovic, and M. Borenovic, “Indoor Localization

Using Multi-operator Public Land Mobile Networks and Support Vector Machine Learning Algorithms," *Wirel. Pers. Commun.*, vol. 104, no. 4, pp. 1573–1597, Feb. 2019, doi: 10.1007/s11277-018-6099-1.

[68] I. S. Al-Mejibli, D. H. Abd, J. K. Alwan, and A. J. Rabash, "Performance Evaluation of Kernels in Support Vector Machine," in *2018 1st Annual International Conference on Information and Sciences (AiCIS)*, IEEE, Nov. 2018, pp. 96–101. doi: 10.1109/AiCIS.2018.00029.

[69] H. Saadatfar, S. Khosravi, J. H. Joloudari, A. Mosavi, and S. Shamshirband, "A New K-Nearest Neighbors Classifier for Big Data Based on Efficient Data Pruning," *Mathematics*, vol. 8, no. 2, p. 286, Feb. 2020, doi: 10.3390/math8020286.

[70] L. Zhang, X. Meng, and C. Fang, "Linear Regression Algorithm against Device Diversity for the WLAN Indoor Localization System," *Wirel. Commun. Mob. Comput.*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/5530396.

[71] S. Zhang, P. Du, C. Chen, W. De Zhong, and A. Alphones, "Robust 3D Indoor VLP System Based on ANN Using Hybrid RSS/PDOA," *IEEE Access*, vol. 7, pp. 47769–47780, 2019, doi: 10.1109/ACCESS.2019.2909761.

[72] X. Cui, J. Yang, J. Li, and C. Wu, "Improved Genetic Algorithm to Optimize the Wi-Fi Indoor Positioning Based on Artificial Neural Network," *IEEE Access*, vol. 8, pp. 74914–74921, 2020, doi: 10.1109/ACCESS.2020.2988322.

[73] Z. Feng, Y. Cao, and J. Yan, "A Received Signal Strength Based Indoor Localization Algorithm Using ELM Technique and Ridge Regression," in *2019 IEEE 2nd International Conference on Electronic Information and Communication Technology (ICEICT)*, IEEE, Jan. 2019, pp. 599–603. doi: 10.1109/ICEICT.2019.8846396.