T.C.
SIIRT UNIVERSITY
INSTITUTE OF SCIENCE

# IMPROVING INDOOR POSITIONING SYSTEM BY USING WI-FI FINGERPRINT WITH MACHINE LEARNING METHODS

**MASTER DEGREE THESIS**

**Student: YADGAR AHMED HASSAN**
**(153111014)**

**Department of Electrical and Electronics Engineering**

**Supervisor: Asst. Prof.Dr. YILMAZ KAYA**

**May - 2017**
**SIIRT**

**THESIS ACCEPTANCE AND APPROVAL**

The thesis titled "IMPROVING INDOOR POSITIONING SYSTEM USING WI-FI FINGERPRINT BY MACHINE LEARNING METHODS" prepared by Yadgar Ahmed HASSAN was unanimously approved by the following jury on 25/05/2017 as a master's thesis in Siirt University Institute of Science and Technology Electrical and Electronics.

| Jury Members | Signature |
|---|---|
| **President**<br>Asst.Prof. Ramazan TEKİN | ………………… |
| **Supervisor**<br>Asst.Prof. Yılmaz KAYA | ………………… |
| **Member**<br>Asst.Prof. Volkan Müjdat TİRYAKİ | ………………… |

I confirm the above results.

Assoc.Prof. KORAY ÖZRENK
Director of Institute of Science

This thesis study ……………. By ………… The project was supported by numbered.

**THESIS NOTIFICATION**

This thesis, which is prepared in accordance with the thesis writing rules, complies with the scientific code of ethics, in case of exploitation of others' works it is referred to in accordance with the scientific norms, innovations contained in the thesis and the results were taken from somewhere else, I declare that any part of the thesis that there is no tampering with the used data is not presented as another thesis work at this university or another university.

Signature

YADGAR A. HASSAN

Note: In this thesis, the use of original and other source notifications, tables, figures and photographs without reference, it is subject to the provisions of Law No. 5846 on Intellectual and Artistic Works.

## ACKNOWLEDGEMENT

Alhamdulili, all praises be to **ALLAH**, the most Beneficent and the most Merciful, the Exalted, the most Gracious, the Victorious, the Compeller, the Knower of All who enabled me to undertake and complete my reserch work and finally write up the outcome report leading towards the fulfillment of Master of Engineering degree.

At the start I shall say " I am deeply indebted to my dad is **Ahmed Hassan (Hajy Ahmed hajy Hassan Shamerany)** and my Mom is **Naima Maarouf** and my **all brothers** and **sisters** for their endless moral support and sacrifice to complete this thesis work.

I wish to express my Profound sense of gratitude to my most respect teacher and thesis guide **Asst. Prof.Dr. YILMAZ KAYA**, Lecturer, Department of Engineering, Siirt University, for his inspiration, constant guidance, valuable suggestions, sympathetic advice and unparalleled encouragement made thoughout the course of study. I am really indebted to him.

I want to convey my special thanks and appreciations to **Tech.Augustian Nnamdi.** in the Cyprus International University.

My sincere gratefulness to all of my teacher of in the department especially in the Siirt University for their co-operation and affectionate treatment thought my research work as well as my study , specially **Asst. Prof.Dr. VOLKAN MÜJDAT TİRYAKİ.**

I would like to offer my thanks to all employee in Ministry of Transport and Communications specially staff of halabja governorate branch to help me to complete my study in the Siirt University. Also Many thanks all teachers in the my old university (BSc level) in the Telecommunication engineering of Technical College of Engineering (TCE ) in the Sulaimani Polytechnic University(SPU) in Sulaimani, Iraq.

Finally, again I grateful to **ALLAH**, the Supreme Authority of the University for this entire reserch work.

Yadgar A. Hassan

SIIRT-2017

# CONTENTS

**Page**

## LIST OF TABLES

**Page**

## LIST OF FIGURES

**Page**

# ABBREVIATIONS AND SYMBOL LISTS

| **Abbreviation** | **Explanation** |
|---|---|
| **WLAN** | : Wireless Local Access Network |
| **IPS** | : Indoor Positioning System |
| **GPS** | : Global Positioning Systems |
| **UWB** | : Ultra WireBand |
| **RFID** | : Radio-Frequency Identification |
| **IP** | : Indoor Positioning |
| **WPS** | : Wireless Positioning System |
| **IEEE** | : Institute of Electrical and Electronics Engineers |
| **USTTI** | : United States Technical Training Institute |
| **GNSS** | : Global Navigation Satellite System |
| **RF** | : Radio-Frequency |
| **RADAR** | : RAdio Direction And Ranging |
| **iOS** | : iPhone Operating System |
| **BLE** | : Bluetooth Low Energy |
| **GSM** | : Global System for Mobile communication |
| **GPM** | : Global Positioning Module |
| **IMS** | : Internet protocol Multimedia Subsystem |
| **LS** | : Local Server |
| **RSS** | : Received Signal Strength |
| **AP** | : Access Point |
| **kNN** | : k-Nearest Neighbor |
| **DB** | : Data Base |
| **LG** | : Life's Good |
| **PDA** | : Personal Digital Assistant |
| **LFDA** | : Local Fisher Discriminant Analysis |
| **PCA** | : Principal Component Analysis |
| **LDA** | : Linear Discriminant Analysis |
| **ML** | : Machine Learning |
| **MAC** | : Media Access Control |
| **RP** | : Reference Point |
| **ANN** | : Artificial Neural Network |
| **MU** | : Mobile User |
| **RSSI** | : Received Signal Strength Indication |
| **SVM** | : Support Vector Machine |
| **ELM** | : Extreme Learning Machine |
| **MLR** | : Multiple Linear Regression |
| **KTT** | : Karush-Kuhn-Tucker |
| **TP** | : True Positive |
| **TN** | : True Negative |
| **FP** | : False Positive |
| **FN** | : False Negative |
| **RMSE** | : Root Mean Square Error |
| **MBE** | : Mean Bias Error |
| $\mathbf{R^2}$ | : Coefficient of Determination |
| **3D** | : Three Dimension |
| **Std** | : Standard Deviation |

| | |
|---|---|
| **SSE** | **:** Streaming SIMD Extensions |
| **ID** | **:** Identity |
| **VC** | **:** Versatile Customer |
| **ID** | **:** Identity |

| <u>**Symbol**</u> | <u>**Explanation**</u> |
|---|---|
| **dBm** | : decibel-milliwatts |
| **S** | : Signal |
| **T** | : Time |
| **r** | : time-invariant |
| $\boldsymbol{\alpha}$ | : multiplicative signal alteration factor |
| $\boldsymbol{\delta}$ | : sensor noise |
| $\boldsymbol{W}$ | : output weights in input layer |
| $\boldsymbol{\beta}$ | : output weights in hidden layer |
| **C** | : output values of the network |
| $\boldsymbol{H}$ | : hidden layer output matrix |
| $\boldsymbol{Y}$ | : decision feature |
| $\boldsymbol{\hat{\beta}}$ | : output weights |
| **e** | : Error |
| **L** | : Lagrange Problem |
| $\boldsymbol{\xi}$ | : weakness variables |
| $\boldsymbol{\mu}$ | : Lagrange parameter |
| $\boldsymbol{\Phi}$ | : A function in SVM |
| $\mathfrak{R}^d$ | : d-dimensional a feature vector |
| $\boldsymbol{\hat{Y}}$ | : Dependent variable |
| **P** | : Polynomial |
| **exp** | : Exponential |
| **sec** | : Second |
| **obs/sec** | **:** open broadcaster software/second |

## ÖZET

### YÜKSEK LİSANS TEZİ

# Wi-Fi PARMAK İZİ KULLANARAK KAPALI KONUMLANDIRMA SİSTEMİNİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE GELİŞTİRİLMESİ

**YADGAR A. HASSAN**

**Siirt Üniversitesi Fen Bilimleri Enstitüsü**
**Elektrik-Elektronik Mühendisliği Anabilim Dalı**

**Danışman : Yrd.Doç.Dr. YILMAZ KAYA**

**2017, 54 Sayfa**

Yeni nesil telefonlar, PDA, netbook'lar vb akıllı cihazlar Wi-Fi şebekelerine daha duyarlı hale getirilmiş ve bizi bu şebekeler üzerinden ağa dâhil eden son teknolojiler ile donatılmıştır. Bu nedenle, çoğu mobil cihaz, ucuz bir ağ altyapısını kullanan çeşitli hizmetleri sunma imkânını getirdi. Bluetooth ve GPS gibi mevcut hizmetler günümüz ihtiyaçlarını karşılayacak kadar yeterli değildir. Bu çalışmada çoğu akıllı telefonda yaygın olarak bulunan Wi-Fi sinyalleri kullanarak makine öğrenmesi yaklaşımlar ile konum tahmini için alternatif bir servis sunulmuştur. Bu tezde Wi-Fi teknolojilerinin konum tahmini için kullanılabilirliğini deneysel çalışmalarla göstermiş olduk. Her biri 4 katlı olan 3 farklı bina için Wi-Fi verilerinin RSS değerlerine dayalı olarak yakınlık tahmini yapan bir model önerdik. Ayrıca tezde ELM, ANN, SVM ve Regresyon gibi makine öğrenmesi yöntemler ile Wi-Fi parmak izi kullanılarak konum belirlemenin gerçekleştirilebileceği gösterilmiştir. Önerilen yaklaşımları test etmek için herkese açık olarak paylaşılan UJIndoorLoc veri seti kullanılmıştır. Önerilen yaklaşımlar ile yüksek tanıma ve başarı oranları elde edilmiştir.

**Anahtar Kelimeler:** İç mekan konumlandırma, kablosuz parmak izi, Sınıflandırma, Makine Öğrenmesi.

# ABSTRACT

**MSc THESIS**

**IMPROVING INDOOR POSITIONING SYSTEM BY USING WI-FI FINGERPRINT WITH MACHINE LEARNING METHODS**

**YADGAR A. HASSAN**

**The Graduate School of Natural and Applied Science of Siirt University**
**The Degree of Master of Science**
**In Electrical-Electronics Engineering**

**Supervisior : Asst. Prof.Dr. YILMAZ KAYA**

**2017, 54 Pages**

As devices (including smartphones, Personal Digital assistants, netbooks etc) are been equipped with state of the art technology making them susceptible to new generation Wi-Fi networks, getting more integrated and getting us more interconnected. Because of this, most mobile devices has brought with it the possibility of providing a lot of various services that utilize a cheap network infrastructure. Existing approaches such as Bluetooth and GPS signal services are not sufficient enough to meet the requirements of accuracy and flexibility. In contrast, Wi-Fi, which is commonly available on most smartphones, provides a compelling alternative for proximity estimation using machine learning tools. In this thesis, we demonstrate through experimental studies the efficacy of Wi-Fi for this exact purpose. We propose a proximity estimation model to determine the distance based on the RSS values of Wi-Fi data in 3 different buildings, each having at up to 4 floors. Thesis has explored a Wi-Fi fingerprint based localization scheme that exploits the power of machine learning methods to solving the problems of Wi-Fi fingerprint localization. The data used was acquired from UJIndoorLoc dataset wish is publicly available and the recognition rate are impressive.

**Keywords:** Indoor localization, Wi-Fi fingerprinting, classification, machine learning.

# 1. INTRODUCTION

The prevalent availability of Wireless Local Access Network (WLAN) and Wireless Fidelity (Wi-Fi) on most mobile devices (including smartphones, Personal Digital Assistants, netbooks etc) has brought with it the possibility of providing a lot of various services that utilize a cheap network infrastructure (Addesso et al. 2013). Indoor Positioning System (IPS) seeks to connect people or objects wirelessly inside a building using magnetic sensor networks, or other sources of data. The main end-user benefit of IPS lays on the extension of mobile computing, which are location location-aware to indoors, such as augmented reality, store navigation, etc (Li, Le, and Wu 2012). Recently, many services are taking advantage of IPS and tracking of mobile device location in such areas as; where there is need to locate individuals in perilous ranges, confinement of companions with regards to social condition and also area focused on data in shopping centers and places of tourism.

However, as more mobile devices become pervasive, context aware applications are becoming an import area of interest for developers, due to the fact that majority of applications presently, rely solely on Global Positioning Systems (GPS) and they tend to function poorly while been operated indoors. At the moment, a de facto standard is yet to be established for the IPS system design. Since the widespread use of both WLAN and mobile devices, Wi-Fi-based Indoor Positioning System has hence come into play and is a valid approach for IPS requiring no extra cost (Ferris, Fox, and Lawrence 2007).

The research presented in this thesis shows how using machine learning is used to improve Wi-Fi fingerprint on indoor positioning system. Currently, there are many technologies used to build robust services based on location-aware application. Global Positioning Systems (GPS) were one of these services in well-known and accurate method used mostly outdoors and does not really operate effectively indoors (Gupta and Sutar 2014).

GPS has is mainly been used in outdoors, it works well in well positioned environment. As the GPS technology majorly depends on the signal propagation via air medium, architectural structures such as buildings and complexes definitely interferes with its signal propagation and hence limits the usability of GPS in indoor environment. In order to fulfil the need of Indoor Positioning (IP), researchers in the field of IP have found alternative ways to the use of existing indoor wireless communication technologies, like Ultra WireBand (UWB), Radio-Frequency IDentification (RFID), Zigbee, Bluetooth and Wi-Fi. Early development of Wi-Fi technology has subsequently lead to the integration and availability of Wi-Fi networks on verities of different devices (Ma et al. 2015) which have built-in Wi-Fi receiving module, hence IP becoming more attractive area of research in order to improve its positioning system.

Due to the increase in usage and implementation of WLAN and mobile devices, Wi-Fi-dependent IPS has become a practical and valid method for IPS that does not require extra hardware or equipment cost. But suffice is to say that, Wi-Fi-based IP system as (WPS) accuracy depends upon the number of wireless position which are stored in the dataset, “whereas the possible signal fluctuations that may occur can increase errors and inaccuracies in the path of the user”(Li, Le, and Wu 2012).

Machine learning is a field of artificial intelligence dealing with algorithms that improve performance over time with experience. Supervised learning algorithms for regression are trained on data with the correct value given along with each variable. This allows the learner to build a model based on the attributes that best fit the correct value. By giving more data to the algorithm the model is able to improve. Learning can be described in this way as improving performance. The measure of performance is how well the algorithm predicts the regression value given a set of variables or attributes.

Finding the building models that generalize well given large amounts of data with many attributes by discovering patterns and trends in the data is a task very difficult or impossible to solve except with the use of Machine learning algorithms (Mascharka and Manley n.d.). Now that sensors are increasingly been made available in most mobile devices, it is evident that large amount of data can be collected and used to aid the features in the localizations process. Machine learning algorithms have therefore become a natural solution for sifting through these large datasets and determining the important pieces of information localization using Wi-Fi fingerprinting, in other to build accurate models that predict an IP.

### 1.1. Wireless Basic

Wi-Fi is an acronym that stands for Wireless Fidelity. It mainly refers to any type of IEEE 802.11 WLAN. By origin, the University of Hawaii, was one of the pioneers of Wireless Local Area Network (WLAN) during the early 1970s when they developed a wireless a network which was used to wirelessly transmit data within the Hawaiian Island. Later on in 1991, the Institute of Electrical and Electronics Engineers (IEEE) started to examine institutionalizing WLAN innovations. In 1997, the IEEE approved the first 802.11 standard which refers to WLAN.

Wireless networks employs most of the components similar to wired networks, however, wireless networks usually needs to convert information signals into a form suitable for transmission through the air medium. (Chapter 1 , Wireless System Architecture : How Wireless Works 2008).

A wireless network is made up of several constituents that bolster interchanges utilizing radio or light waves that is proliferated through the air medium. Some of these components cover with those of wired systems, yet uncommon thought is important for these parts while sending a remote system (Chapter 1 , Wireless System Architecture : How Wireless Works 2008).

### 1.2. Wireless Security

There are so many issues bordering around wireless communication and these can be linked to the security of information confidentiality. This is caused by the fact that data from wireless networks which are transmitted via radio waves are vulnerable to interception from unauthorized persons. The secure use of wireless networks is based on users connected to the network via predetermined access points using protocols in order to access the network securely (Ijeh et al. 2009). To discuss any issue related to Wi-Fi technology, security of information been transmitted is paramount. It is necessary to explore recent advances in wireless security in order to pin point security issues related with trust, management, interoperation and measurement. In order to solve the potential problems of wireless network security in the contest of this thesis, the problems must be treated independently (Evans, Wang, and Ewy 2006).

For IPS systems developed to make use of the ad hoc remote system it is accept that no pre-conveyed framework is accessible for directing bundles end-to-end in a system, rather it depends on middle person peers. Securing specially appointed steering presents challenges in light of the fact that every client conveys to the system their own particular versatile unit, without the incorporated arrangement or control of a customary system. (Sanzgiri et al. 2002) has it that if an efficiently secured routing algorithm prevents exploits, it must at least ensure that no node can prevent successful route discovery and maintenance between any except the non-participating nodes.

### 1.3. Motivations of The Study

With the advent of smartphones, GPS has been a very key components that GPS ships it's sensors with. Hence the interest of researchers, government, companies and consumers at large in providing location services for user and devices. Governments in corporations with companies have made enormous investments in creating infrastructure for map based-location services (GPS Use in U.S. Critical Infrastructure and Emergency Communications Presented to the United States Technical Training Institute ( USTTI ) n.d.). Global Navigation Satellite System (GNSS) Asia reports that "over the coming decade the introduced base of GNSS gadgets will increment right around four-crease. It is normal that the quantity of GNSS gadgets will increment in Europe and North America from 1 to 3 for every occupant over the coming" decade (Ugave 2014). Smartphones are going to dominate the global GNSS revenues and are also expanding into other market segments.

Another motivation to this thesis is the fact that it is quite tedious to find special places like stairways, conference rooms, restaurants, store navigation etc. and very demanding. Sometimes these buildings are inhabited by people with special needs, who need not worry too much about navigating around their environment (Li, Le, and Wu 2012). Also, it is not very easy for some people to information signs, such as the blind and some people who are new to a place or are in a rush. Indoor localization systems is mostly used in context-based targeted advertising, emergency response and assisted living, robotics applications, and indoor navigation in places such as airports, malls, and campuses. Again, in addition, we must learn the properties of a building before we can start searching inside it to find the kind of things we need. In the university campus environment, finding classrooms, faculties, seminar or conference halls and some other even is a usual daily activity. With the advent of indoor positioning system, meetings can be arranged for students, and faculty members.

Events locations can be loaded to the map, this way there is no need for extra activities and time consuming announcements. If there is any emergency situation, people can be directed through the emergency exits or "administration to safe spots.

This is a standout amongst the most essential parts of position-based administrations since it can spare lives. For instance if there should be an occurrence of flame application can control individuals to outside stairways, or if there should arise an occurrence of quake, it can direct you to the sheltered spots".

**1.4. Indoor Localization**

Indoor localization is of immense significance for a wide range of mission-critical applications, "which have pulled in a great deal of research exertion in the previous decades. Most radio-based arrangements require a procedure of site overview, in which radio marks of an intrigued zone are commented on with their genuine recorded areas. Site study includes serious expenses on labor and time, restricting the pertinent structures of remote confinement around the world (Chen, Yenamandra, and Srinivasan 2015). Most works "in the field of indoor localization included the adjustment of standards of RADAR based target finding to the RF based approach in indoor conditions. RADAR" (Bahl and Padmanabhan 2000) "was one of the spearheading works in the field of indoor remote restriction in which the authors confined clients inside" a building and this study has since a hot area of study by many researchers alternatively, researchers have studied using Wi-Fi transmissions for localization and activity/gesture recognition. This thesis differs from other lines of work as we seek to use machine learning to localize the Wi-Fi fingerprint.

**1.5. Problem Statment**

UWB, RFID, Zigbee and Bluetooth are limited in range. GPS on the other hand works excellently outdoors but, has a lot of limitations in estimating the position of devices indoors. Just like position-based services used by Google maps is the GPS for discovering the position of a devices. In creating an IPS, we will be faced with a problem because GPS is not efficient indoors or does not work at all. Looking more closely, we discover another problem which is the fact that services like Google maps does not have the maps of all the building or cannot locate the position of objects in the building. Without having a precise map, navigation and positioning becomes useless. Research into indoor positioning has brought about algorithms that answer the question to the problem is different dimensions.

Wi-Fi network infrastructure is found in many public facilities such as campus, shopping malls, airport etc. and therefore can be used for indoor positioning. In addition, the perversity of Wi-Fi-capable devices makes this approach especially cost effective although the signal propagation model of Wi-Fi fingerprint depends on the signal arrival angle, the delay and the signal power.

### 1.6. Literature Summary

The thesis attempts to use machine learning methods to provide a more robust IPS using available Wi-Fi fingerprints. This is due to the prevalence of Wi-Fi and its integration and penetration to most mobile devices and ubiquitous systems. GPS is an already existing and efficient positioning system, but it does not really work in doors and so there is need to use Wi-Fi fingerprints with the help of machine learning methods we can locate places within a building. This is very important when it comes to helping people navigate their way in massive building. It is of immense advantage to elderly people, disabled and people with special care. Emergency exit ways can be suggested to users of a massive building during rescue or emergency situations.

### 1.7. Advantages of IP System

The use of IPS in various real world applications is widening and the advantages are quite numerous to mention some of which are mentioned herein;

1. **Accuracy**: Accuracy of an IPS is a very important element in design and implementation of the entire system. Using GPS as for positioning in an outdoor environment has gained a lot of research and implementation but it is not very accurate when compared with using Wi-Fi in an indoor environment. According to (Jiang et al. 2015) and (Torres-Sospedra et al. 2015) high accuracy has been obtained in an indoor system using Wi-Fi alongside with various localization algorithm.
2. **Versatility**: IP Systems are usually versatile in the sense that, if a system using sensor emitting devices is switched off, the systems are still able to adapt to new changing environment and behave as if the new device has been there all the while.
3. **Minimal infrastructure:** IP systems are designed to make use of existing WiFi in a building and integrated using existing sensor devices on the mobile devices which have become predominant.

Most IPS systems works interchangeably with either WiFi or Bluetooth or a combination of both, making user able to use existing hardware that is available and cutting the cost of purchasing new hardware.

4. **Platform robustness**: Also, developers can deploy application on user devices through the android or IOS platforms and provide updates as frequent as new algorithms are implemented on the system. This provides a robust way of providing the services to variety of users (Husen and Lee 2016).

5. **Orientation and Position:** Most IP systems are able to estimate the exact path of user following the user that carries a smartphone and includes the orientation at every instant without the detected person having idea.

### 1.8. Disadvantages of IPS

Although IP system is a very important part of localized positioning, it also has some disadvantages as highlighted herein.

1. The accuracy measure is within 5-15m compared to using RFID which is less.
2. Incompatibility with some operating systems like iOS is an issue. This is due to the fact that WiFi client based positioning system is not yet possible with the iOS devices but the use of BLE can be an alternative on iOS instead of WiFi.
3. Application is required IP systems come with applications that need to be developed, but some users may not have access to these application or able to pay for them since they not free application.

### 1.9. Thesis Outline

This thesis is organized into five chapters. Chapter 1 introduces the whole work and provides and overview of the subjects discussed herein. Chapter 2 provides a general summary of related research in the field of indoor localization. Chapter 3 presents the main work done in this research, the steps taken and general the work will be examined completely including methodology used. Chapter 4, presents the results of the experiment of algorithm that we have used are illustrated. Lastly in Chapter 5 conclusion is made and we discuss future work.

## 2. LITERATURE REVIEW

Location based services are of immense advantage in several real world applications, hence, no surprise about lots of research that have been made as regards getting to this point in the development of IPS systems and ongoing researches as well.

### 2.1. Categories of Methods Used

There two major methodologies available for indoor positioning system (IPS), the first one is using a special hardware for this purpose which has been demonstrated a high accuracy. Some research made use of RFID in (Thiem et al. 2008) and (States 2004), infrared (Kimura 2012) and (Weng et al. 2009), or ultrasound (Piontek and Seyffer 2007), (Lee 2014) and (Lemiux and Lutfiyya 2009) that has been successful with a high accuracy, but installation of the hardware is needed which is not cost-effective (Saeid 2013).

The second method used in indoor positioning is by the use of already installed hardware like Wi-Fi infrastructures (Rashid, Chowdhury, and Nawal 2016),(Ma et al. 2015), GSM signals (Ugave 2014), and also Bluetooth signals (Bekkelien 2012) in buildings by measurement of received radio signals. This data is captured and pass through the tedious training process and classified using unique machine learning methods which are tested in this thesis.

### 2.2. Review of RFID in IPS

Most times, RFID systems are referred to as merely glorified bar-code systems. This is a riskily and limiting approach that could make someone think of it as a very small subsection of the anticipated benefits of RFID. Trying to Compare RFID systems with bar-code systems is useful for understanding how these systems differ and what challenges they face.

The use of RFID “frameworks are as of now regular in numerous business conditions, for example, shopping centers, markets, coordinations, cargo administrations and support. Verification and confinement of people is as yet not exceptionally prominent but rather could be extremely useful for the individual itself and in addition for administrators of e.g. air terminals, shopping centers or public expos. Upgrading the adaptability of RFID based frameworks by utilizing remote advances for systems administration obviously enhanced the ease of use and acknowledgment of RFID arrangements. In this manner, an approach for a (minimal effort) RFID peruser arrange exclusively in light of remote interchanges is proposed“. (Thiem et al. 2008) describes their “present and continuous advancements of remote associated RFID per-user frameworks utilized for assessment of an aloof client limitation framework. In their review, a two-level design of various remote associated RFID per-users was presented. In the lower level, i.e. the main level, a few remote advancements are worked to can exist together utilizing various types of correspondence advances to associate distinctive RFID perusers. These advancements are associated through a completely fit WiFi/WLAN system to bring them into the IP world”. They are yet to report performance, but the research is positive to produce high performance.

### 2.3. Blutooth IPS

In a research conducted by (Bekkelien 2012), implemented an indoor situating framework depending on Bluetooth innovation, and incorporated it into the Worldwide Situating Module (GPM) which was created at the Organization of Administrations Science at the College of Geneva. This review looks at the flag qualities of encompassing Bluetooth gadgets estimations put away in database taken over the indoor territory, in order to appraise the client's position. Through an assessment of the framework, an exactness of roughly 1.5 meters has been acquired.

### 2.4. WI-FI Fingerprinting Utilizing Machine Learning Methods

Many existing services and even upcoming administrations may exploit the learning and following of a Wi-Fi fingerprinting from mobile nodes or where ever in order to localize “individuals in unsafe ranges, limitation of companions in a social situation, or area focused on data in strip malls or touristic places”. The remaining part of this chapter attempts to review related works in the area of indoor positioning, Wi-Fi fingerprinting enhanced with Machine learning capabilities.

In (Addesso et al. 2013) WLAN 802.11x technology was used to create an indoor vehicle location system which as shown by the study, exploits “an as of now introduced foundation. While the WLAN 802.11x innovation” powers the “proposed engineering, the administration rationale, which permits to offer a few administrations to every vehicle utilized amid test and furnished with a Wi-Fi empowered gadget, is given by a nearness benefit coordinated in IMS. The guess and seeing of a vehicle's position inside the indoor stopping range is completed by a modified Area Server (LS), additionally as indicated by fingerprinting calculations in view of the RSS (Got Flag Quality) limits of signal edges that client customers constantly get from the Get to Focuses (APs).” In the methodology of this Study, majority of algorithms conventionally makes use of localization and exploits the triangulation principle. The positioning algorithms used had “recourse to scene analysis in order to improve precision in the measurements of position for the example, the fingerprinting positioning method based on RSS (Received Signal Strength) is commonly used in Scene analysis. This experiment made use of weighted-kNN algorithm which considers the RSS mean values that are computed for every reference position originally stored in a DB mapper. Secondly, a sequential Bayesian filter was used to regularize the sequence of position estimates by employing a suitable kinematic model, implementing a Kalman Filter in where the estimated position obtained by the weighted-kNN plays the role of the observation”, their classification result was quit improved by the use of the later.

### 2.5. Hybrid IP System Using WI-FI and GSM on ML

Some other researchers have proved in their study that combination of Wi-Fi with GSM RSS data measurement greatly decreases the location estimation error. In this experiment studied by (Bacak and Çelebi 2014), a study was made to estimate different location approaches in order to determine the position of a mobile client indoors. This study shows that “with a specific end goal to decide the position of the versatile client at the trial site (a shopping center), the got flag quality (RSS) RF unique mark based methodologies is considered. Wi-Fi and GSM RSS information were gathered by utilizing LG Nexus 4 PDA at Gebze Center shopping center in Gebze, Kocaeli. The gathered information were prepared utilizing diverse machine learning calculations, number of estimation matrix, number of preparing information, and flag sort on the execution of the limitation frameworks are contemplated and this support the” already stated result.

(Bekkelien 2012) performed a study whose objective “was to execute an IPS contingent upon Bluetooth development, and to consolidate it into the Overall Arranging Module (GPM) made at the Foundation of Organizations Science at the School of Geneva. In trial system, he took a gander at the banner characteristics of incorporating Bluetooth contraptions to a database of estimations assumed control over the indoor locale, remembering the true objective to gage the customer's position. Through an evaluation of the structure, an exactness of approximately 1.5 meters has been gained”. The measured algorithms were, k-NN (1.62 accuracy), k-NN regression (1.60 accuracy) and Naive Bayes (2.13 accuracy).

(Berkvens, Weyn, and Peremans 2015) submit a probabilistic Wi-Fi sensor model which was used for that research to quantify the localization performance of exteroceptive sensors. The calculation performed in this experiment showed the mean common information for this sensor model that they have used. Then, they applied Maximum Probability Estimation and k Nearest Neighbor as a localization scheme to their sensor model and obtain an average Sample Error of 22.15 and 21.97, respectively, on the evaluation dataset.

However, Feature extraction methods have been used by many researchers to extract location features used for indoor positioning in the context of Wi-Fi fingerprinting. But a study by (Deng, Xu, and Chen 2013) made us of enhanced Local Fisher Discriminant Analysis (LFDA) unlike its counterpart linear discriminant analysis and principal component analysis, which suffers from the multimodal property of signal distribution. Their proposal was based on using first of all using LFDA to extract discriminative location features and as such maximize “between-class” distinguishability while preserving “within-class” local structure of captured signal space, which further guarantees maximal discriminative information meant for situating. “At that point, the speculation capacity of LFDA is additionally improved utilizing signal irritation, which creates more number of delegate preparing tests. Test brings about practical indoor condition demonstrate that, contrasted and past element extraction techniques, the proposed strategy diminishes the mean and standard deviation of setting blunder by 23.9% and 33.0%, separately”.

“Fingerprinting situating can be considered as a multiclass order issue. Each reference point can be marked as one class by related RSS tests. The entire procedure of area estimation can be viewed as arranging the ongoing RSS tests into the correct classes. In this way, contrasted and PCA, LDA performs better, since more discriminative elements can be extricated. LDA augments the between-class remove while compelling the inside class separate into a specific degree. At the point when RSS appropriations are Gaussian or semi-Gaussian, LDA performs well” (Koutroumbas, 2008). In any case, in a reasonable remote engendering condition, non-Gaussian and multimodal appropriations can be constantly watched. For instance, because of the radio wire assorted qualities (Bratus et al. 2008) or multipath effect, RSS distribution from one AP at a fixed location may show more than one peak or cluster .

(Thiem et al. 2008) “RFID frameworks are now utilized as a part of many spots and business situations like coordinations, cargo administrations and upkeep. Validation and limitation of people is as yet not extremely prevalent but rather could be exceptionally useful for the individual itself and in addition for administrators of e.g. airplane terminals, shopping centers or expos. Upgrading the adaptability of RFID based frameworks by utilizing remote advancements for systems administration will obviously enhance the ease of use and acknowledgment of RFID arrangements. Along these lines, an approach for a (minimal effort) RFID peruser arrange exclusively in view of remote correspondences is proposed. This approach will limit the establishment overhead and will bring about an adaptable and exceptionally versatile and compact framework with limited administration overhead.”

The “situating of individuals and gadgets precisely both inside and outside has for quite some time been an objective inside the exploration group. The capacity to track focuses of intrigue empowers the acknowledgment of an extensive variety of both setting mindful and calculated applications, which can possibly ease life for a large number of individuals”. There are several applications in the domain of indoor localization using Wi-Fi fingerprinting such as the one depicted in precious paragraph above; vehicle monitoring using Wi-Fi fingerprinting and machine learning methods, other applications are proposed for both the general subject and in addition for work areas, for example, doctor's facilities, portraying the scope of chances for usage of IP algorithms (Mathisen et al. n.d.).

In an experiment conducted by Mathisen et al, a relative assessment of a few indoor situating strategies set at an extensive college healing facility spreading over 160,000m2 — a domain where a few area and setting mindful calculated applications are in day by day utilize, was finished. The situating techniques use estimations of flag qualities utilizing existing Wi-Fi foundation, which facilitates organization and upkeep. They recognized suggestive key measurements which characterize distinctive parts of the strategies' execution. Utilizing these measurements, we moreover give an account of encounters with executing and using indoor situating arrangements in a very assorted condition, in which building sorts and materials, and additionally constructing use vary over the complex. Correspondingly, the assessment information we utilize is assembled at various building complex parts, days, and daytimes, and both at static areas and also going inside the building complex. Our outcomes outline and measure the difficulties and breakdowns in exchanging execution comes about because of a little controlled setting, for example, a little office condition, to an extensive element building complex.

### 2.6. Related Study on WI-FI Fingerprinting Using ML

As technology grows rapidly, new assisted living applications will be created sooner rather than later having client situating as ground innovation: elderly tele-mind, vitality utilization, security and so forth are unequivocally in view of indoor situating data. In a study conducted by (Living et al. 2017) “an indoor positioning system for wearable devices based on Wi-Fi fingerprints was presented. “Savvy wearable gadgets are utilized to obtain the Wi-Fi quality signs of the encompassing Remote Get to Focuses used to assemble a cooperation of Machine Learning arrangement calculations. Once constructed, the outfit calculation is utilized to find clients in view of the Wi-Fi quality signs way gave by the wearable gadget. Their trial comes about for five distinctive urban pads were accounted for, which demonstrated that the framework is powerful and sufficiently solid for finding a client at room level into his/her home. Another fascinating normal for the exhibited framework is that it doesn't require sending of any foundation, and it is subtle, the main gadget required for it to work is a savvy”.

(Ugave 2014) in a study, performed an indoor localization of Wi-Fi fingerprinting using machine learning methods. In this study, they evaluated their systems on a couple of indoor circumstances with arranged characteristics that reflect overhauls more than a couple best in class techniques coordinated from before work. The wide usage of sensors and Wi-Fi compasses can deplete the phone battery in this way they quantitatively spoke to each one of the modules that use the battery control. This audit in like manner performed imperativeness and precision tradeoff examination to give a more broad understanding of how to keenly use these strategies. Furthermore, they investigated, implemented and tested both sensor and machine learning based techniques. With the techniques tested, they achieved an average accuracy between 1-3 meters across most of our evaluated indoor paths.

In this experiment carried out by (RADAR_in_building_RF.pdf n.d.), they presented the use of RADAR, systems on a couple of indoor circumstances with arranged characteristics that reflect overhauls more than a couple best in class techniques coordinated from before work. The wide usage of sensors and Wi-Fi compasses can deplete the phone battery in this way they quantitatively spoke to each one of the modules that use the battery control. This audit in like manner performed imperativeness and precision tradeoff examination to give a more broad understanding of how to keenly use these strategies.

## 2.7. 802.11 L2 Attacks and Offensive Fingerprinting

In discussing about IP using Wi-Fi fingerprinting, there is need to also talk about some of the security vulnerabilities that may occur. According to (Bratus et al. 2008), several vulnerabilities have been discovered recently in diverse implementations of the 802.11MAC layer, which resulted in increased interest in fingerprinting of 802.11 phases for both antagonistic and defensive occupations. The Blackhat 2006 open appearing of a section level experience for a 802.11 card driver has affirmed figures that the versatile nature of 802.11 traditions suggested exploitable vulnerabilities and made basic measures of introduction. This release was joined by reports of productive fingerprinting of hardware programming blends, in light of their association layer direct, as a strategies for driving completely centered around ambushes (Arackaparambil et al. 2010).

## 3. MATERIALS AND METHODS

Indoor positioning using Wi-Fi fingerprinting and the different machine learning algorithms applicable. Indoor radio banner multiplication is difficult to suspect because of number components, which cause the obscuring of the radio banner. This obscuring makes the radio banner vacillate. As this survey focuses on change of machine learning based IPS (Box and Luang 2010). There are two phases in the Wi-Fi fingerprinting-based techniques; viz, offline, in which case, the recognized RSS of the Wi-Fi from different Aps will be recorded at alluded to ranges, which are known as reference centers (RPs), then later saved in the database (Fingerprint-Based Technique for Indoor Positioning System Via Machine Learning and Convex Optimization 2016), for the online query to take place and in the online methods, the reported RSS of ; these will be evaluated in this chapter (Liu, Jiang, and Striegel 2014).

Using RSS, (Kupper, 2005). Backpropagation learning figuring, which is a Artificial Neural Network (ANN) based estimation, is utilized as a part of this audit for indoor arranging with the intent to assess the zone of the MU with irrelevant slip-up. The clarification behind picking machine learning systems is a consequence of its healthiness against bustle and deterrent which are one of the fundamental contemplations affecting the precision of IPS. Figure 3.1 below shows the structure of the system.

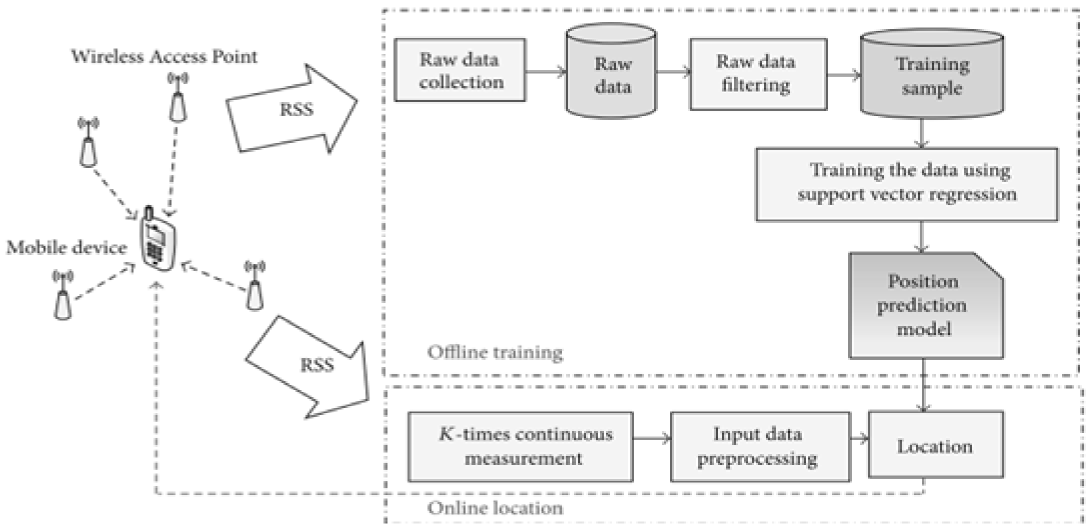


**Figure 3.1.** Structure of the system(Ma et al., 2015)

### 3.1. Materials

Torres and his colleagues at work (2014) the shared data set will be used. From a literary perspective, it is seen that there is no common database while there are studies for IP systems. UJIIndoorLoc database, it is presented to fill this gap. This dataset contains three buildings from Universitat Jaume with a total area of about 110.000m2 from a total of 4 or more floors. This data set can be used for classification problems for floor or building estimation, a dataset that can be used in regression problems for latitude and longitude estimation. The data set was collected in 2013 with 20 different users and 25 different Android devices. The data set consists of 19937 training and 1111 test records and 529 attributes. The data set includes fingerprints of wireless devices, coordinates from which the signals are received and it consists of other information. Each fingerprint is expressed in terms of wireless access point and received signal strength. In the data set, these signal intensities are expressed as negative values between -104 dBm and 0 dBm. The values of 100 in the data set are used when no signal strength is detected. When creating the data set, 520 different access points were detected.

### 3.2. WI-FI Based Positioning Systems

Wi-Fi localization fingerprint has become one the most widely used techniques in terms of indoor positioning; it has two distinct phases which include; 1st phase, ususally called the offline phase, in which the RSS from all the AP is gathered from different locations within a building such that location matrix of the data is generated from this operation giving the fingerprint of each and every AP in that location. The second phase is called the online phase, in which several pattern matching algorithm is used to comprehend the range grid and find out the zone of Versatile Customer (VC) by differentiating the RSS being seen by the MU with the one recorded in the zone structure (Box and Luang 2010). Varieties of existing methods have been proposed to find the closest match in the fingerprints, which include *k*-nearest neighbor (KNN) (Zandbergen 2012), probabilistic method, neural networks (Box and Luang 2010), and decision tree. In KNN method, fingerprints are collected by deploying the mobile devices at each predefined reference point in the operation area. “Then $k$-nearest position is found based on the least squares method during the online positioning phase. At last, the location of the target is determined as the average of these k-nearest positions. To improve the accuracy of the KNN positioning method, the users’ feedback during the online location period is considered. The Smallest $m$-vertex polygon method is similar to the KNN method”.

### 3.3. Data Collection/Offline Module

The quality of the training sample determines the quality of the final position prediction model when using the support vector regression algorithm: the better the quality of the training sample is, the better the final model is. Meanwhile, the more precise the input of the model is, the more accurate the location results obtained are. In the wireless indoor environment, the wireless signal transmission is influenced by many factors, and the RSS information received by the mobile device is variable due to complex transmission environments (Husen and Lee 2016). Even at a fixed point, the RSS information collected by the mobile device will be changing, which may compromise the quality of the training sample and the model input and affect the final position accuracy. “Therefore, in order to obtain the high quality training samples, the raw data needs to be filtered at the offline phase. And the input RSS information must also be preprocessed to obtain a more accurate input at the online phase. Therefore, we will discuss the raw data filtering rules in this section and the input RSS information preprocessing in the next section. There are a lot of existing methods to ease the effects of abnormal data. For example, averaging is a simple and common used method, which can give a measure that is more robust in the presence of outlier values. Raw data filtering can be done through the simple averaging of the RSSI value groups. For wireless environment, reliability is always a concern, which makes the packets containing RSS information vulnerable to loss and leads to lower RSS value. Based on this intuition, the rule of removing the least values eliminates certain number of the lowest values from a RSSI value group”. However, 802.11 radio transmissions have their own characteristics. To get better result, filtering rules should be designed according to these characteristics (Shi et al. 2015).

### 3.4. Selecting WI-FI Fingerprint

While working in the offline stage, the key to formulate a stable location fingerprint is selecting the best appropriate signals from the database can be gathered, by making effort to remove useless signals. However, their approach still lacks in producing highly stable location fingerprints because the authors definition and identification methods of useless signals is not enough to select the best signals to form a good location fingerprints (Husen and Lee 2016). An earlier work focusing on orientation-based indoor localization using Wi-Fi fingerprint, the study managed to localize the Wi-Fi by location and orientation. However, there are still rooms for improvements in the calibration phase.

In this thesis, a novel approach to formulate a highly stable indoor location fingerprints by eliminating offline recalibration to produce an accurate and robust indoor localization system is introduced.

### 3.5. WI-FI Received Signal Strength Model

“The *m*-dimensional Received Signal Strength (RSS) vector, $\{s_i(t)\}_j$, at time *t* at a particular calibration location *j* due to *m* Wi- Fi sources, i=1, …, m, is represented as

$$\{S_i(t)\}_j = \{s_{1,j}(t),\ldots,S_{2,j}(t),\ldots S_{i,j}(t),\ldots,S_{m,j}(t)\} \tag{3.1}$$

From the representation above, the bold-face letter, **s**, is to represent it as a random variable.

Here, we model $\mathbf{s}_{i,j}(t)$, the RSS from the ith Wi-Fi source at the jth calibration location, as

$$\mathbf{S}i,j(t) = \boldsymbol{\alpha}i,j(t) \times r_{i,j} + \boldsymbol{\delta}_{i,j} \tag{3.2}$$

Where $r_{i,j}$ represents the time-invariant RSS with no spatiotemporal disturbances present, $\alpha_{i,j}(t)$ the multiplicative signal alteration factor due to the spatiotemporal disturbances of $r_{i,j}$, and $\delta_{i,j}$ the sensor noise. It is noteworthy that we introduce $r_{i,j}$ as the ideal time-invariant signal attenuated by the distance from the Wi-Fi signal source i to the region j through the invariant channel traits, taking simply settled building establishment and furniture outline into thought. The actual signal, $s_{i,j}(t)$, is then considered as the alteration of $r_{i,j}$ by $\alpha_{i,j}(t)$, $0 < \alpha_{i,j} \leq 1$, reflecting the stochastic channel properties on account of subjectively moving people or possibly dissents and moreover as a result of discretionary presentation of the PDA customer which ruins the channel in the midst of RSS estimation. Observe that $s_{i,j}(t)$ with minimal random spatiotemporal disturbances implies $\alpha_{i,j}$ to be equal or close to 1".

### 3.6. Online Module

"During online location phase, mobile devices collect the RSS information from APs (Access Points) and use this information as the input to the location model established in the offline training period to get the real location. The process is shown as follows".

(i) *k-Times Continuous Measurement to Collect RSS Information*. The RSS information collected by the mobile device is vulnerable, changeable, and inconsistent even at the same position. To get the accurate RSS information, mobile devices execute *k* continuous measurement and get *k* RSS values at each location during online phase.

(ii) "*Input Data Preprocessing.* Input data preprocessing analyzes *k* RSS values obtained through *k*-times continuous measurement, eliminates the effects of environmental disturbance on the RSS values, and determines the final RSS value that describes the relationships more accurately.

(iii) *Location*. The RSS information after preprocessing is taken as the input to the position prediction model trained during offline phase and the final location is calculated.

In this stage, mobile devices do not need to exchange the information with the central server. This is because the SVM model in this study is small enough to be stored in the mobile device". And unlike the common fingerprints based methods, there is no need to search all the reference points to match the location.

"Offline data filtering and online data preprocessing can effectively reduce the impact of transient variations of RSS values on the location accuracy. However, when major or permanent changes happen, for example, the location of AP changes or the indoor layout changes, the position prediction model needs to be retrained by the latest collected data. This retrained process can be triggered by the location error analysis. If the location error exceeds predefined threshold, the retraining begins afresh”.

### 3.7. Machine Learning Methods/Techniques

The data in this study will be analyzed with different machine learning methods. Artificial neural network (ANN), support vector machines (SVM), extreme learning machines (ELM) and regression methods will be used. All algorithms will be performed according to a 10-fold cross – validation test.

#### 3.7.1 Artificial neural network

ANN is computer systems that are automatically developed without any help, such as the ability to derive new information, create and discover new information through learning from the characteristics of the human brain (Kaya et al., 2014). ANNs have been set up to overcome the hindrances of the customary approaches to manage handle complex issues. This system picks up from given cases by building up an information yield mapping in order to perform desires. The ANN models work like a black box without requiring the point by point information of a structure. As opposed to requiring this information, they take in the association between the data parameters and the controlled and uncontrolled elements by focus the previously recorded data like non-direct backslide. One more great position of using ANNs is the capacity of directing broad and complex structures with a boundless number of interrelated parameters. An ANN model is made out of an information layer, no less than one covered layers, and a yield layer. All things considered, the amount of neurons in the data layer is compared to the data number in the issue while the amount of neurons in the yield layer is contrasted with the pined for yield number. The amount of hid layers and the neuron number in the covered layers are controlled by trials. An ANN model is shown in Figure 3.2.One of the most important elements of an ANN is the connections that nerve cells transmit data to each other. A connection that transmits information from one cell to another cell has a weight value.

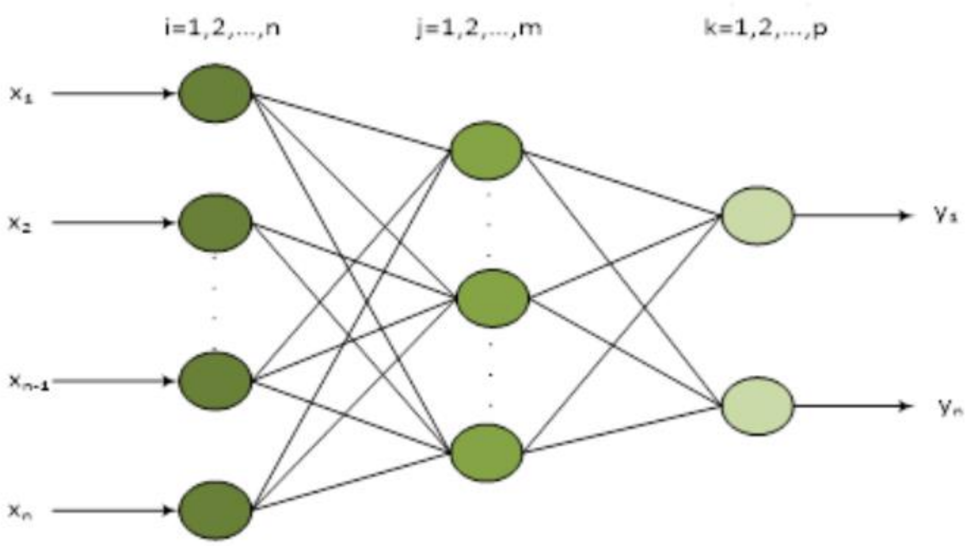


**Figure 3.2.** Simple one ANN architecture(kaya and Kayci, 2014).

The nerve cells in each of these three layers and the weights connecting them are shown in Fig. 3.3. In Fig. 3.3, the circles show the nerve cells and the lines connecting the cells show the weights. One of the most important elements of an artificial neural network are the connections that nerve cells transmit data to each other with these connections. G (x) in Fig. 3.2. is the summation function, which calculates the net input to a nerve cell. By multiplying the input variables by the weight coefficients, an input is generated for the G (x) summation function. A nerve cell structure is given in Fig. 3.3.

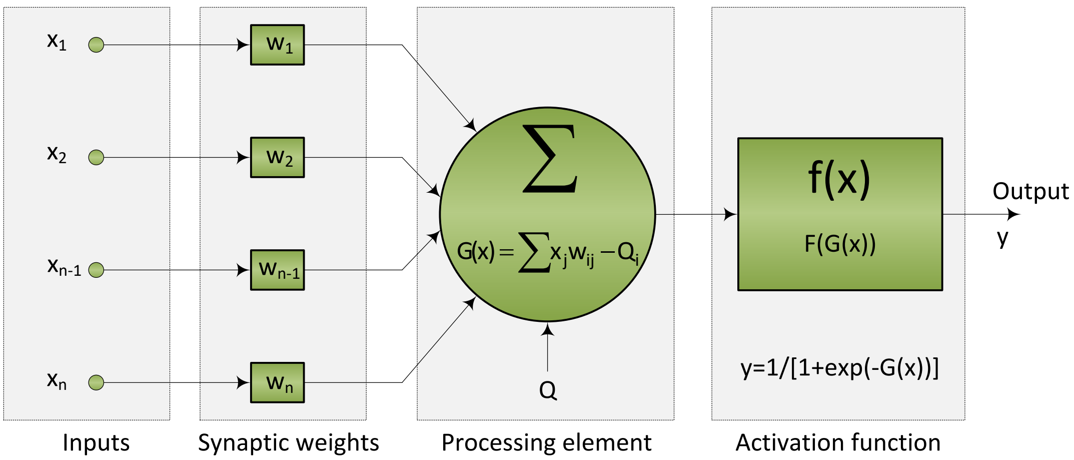


**Figure 3.3.** The model of an artificial neuron(kaya and Kayci, 2014).

Mathematical statement of an artificial neuron can be written as:

$$y_i = F(G(x)) = F(\sum_{i=1}^{n} w_{ij} x_j - Q_i);\ x_i = (x_1, x_2, .......x_n) \tag{3.3}$$

In condition (3.3), ***x={x1, x2, x3, . . . xn}*** is an information variable to be readied. Of course, ***w={w00, w01, . . . wij}*** is the weight and it exhibits the hugeness of data drawing closer to a neuron and the impact on the neuron (Kaya et al., 2007). The estimations of the weights can change at the path toward get ready: *Qi* addresses the breaking point regard; ***F(.)*** is the start work; ***G(.),*** that comes to F(.), is the limit that conveys the yield by taking care of the data. There are particular sanctioning limits, for instance, sigmoid, straying sigmoid, sine and winding reason. In an ANN, most of the neurons may have the same or unmistakable start limits. All together for any ANN models to make the estimations less requesting, establishment limits are required of which derivates can be taken. Which start ability to be used is picked subsequently of the customer trials.

### 3.7.2. Extreme learning machine

Uncommon learning machine (ELM) is a single hid layer feedforward ANN demonstrate whose data weights are delivered erratically and yields of which are handled logically. In the covered layer of ELM, other than having start limits, for instance, sigmodial, sine, Guassian and hard-confine, not in the slightest degree like ANN, non-differentiable or discrete order limits can be used (Suresh et al., 2010).

The execution of standard feedforward ANN depends on upon a couple of parameters, for instance, vitality, learning rate, et cetera. In these sorts of frameworks, parameters, for instance, weights and inclination qualities are required to be revived through edge based learning counts. Regardless, with the target of better learning execution, the arrangement of get ready takes so much time that the learning speed is to an extraordinary degree direct and it is definitely not hard to fall into neighborhood optima. In spite of the way that adding a vitality parameter to the weight modification can cut down the threat of the framework being gotten in neighborhood optima, the time spent on the arrangement method is not lessened. In ELM, input weights and slant qualities are made self-assertively, however yields are gained efficiently (Huang et al., 2006). ELM network is a customized single hidden layer feedforward ANN model, as shown in Figure 3.4.

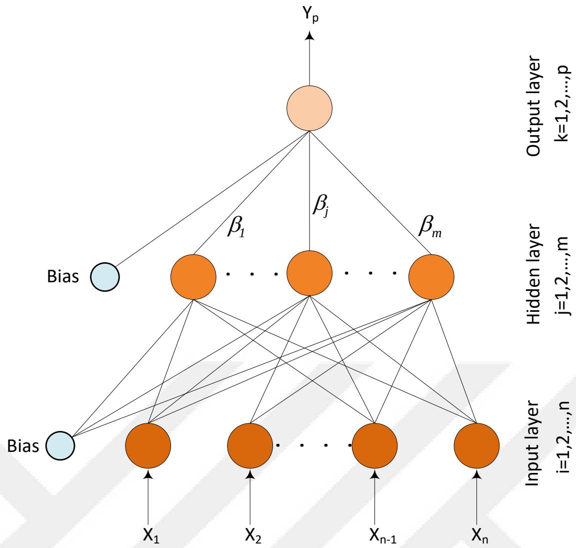


**Figure 3.4.**. A customized single hidden layer feedforward network(kaya et all, 2014).

In here, $X = (X_1, X_2, X_3.....X_n)$ identifies input, $Y = (Y_1, Y_2, Y_3.....Y_p)$ shows the output and network having *m* neuron in hidden layer is stated mathematically as follows; (Suresh et al., 2010).

$$\sum_{i=1}^{m} \beta_i g(W_i X_k + b_i) = O_k \quad , \quad k = 1,2,3.......n \tag{3.4.}$$

where $W_i = (W_{i1}, W_{i2}, W_{i3}.....W_{in})$ indicates output weights in input layer, $\beta_i = (\beta_{i1}, \beta_{i2}, \beta_{i3}.....\beta_{im})$ output weights in hidden layer, $b_i$ threshold values of the neurons in hidden layer and c output values of the network. $g(.)$ is the activation function (Hai et al., 2008).

The purpose in a network with *n* input feature is to provide minimum error. Therefore, Equation (3.4) can be written as follows (Huang et al., 2006).

$$\sum_{i=1}^{m} \beta_i g(W_i X_k + b_i) = Y_k \quad , \quad k = 1,2,3.......n \tag{3.5.}$$

In above mentioned equation, it is possible to write as follows;

$$H\beta = Y \tag{3.6.}$$

where *H, β* and *Y* can be expressed as follows (Huang et al., 2006).

$$H(W_1,........W_M; b_1,........b_m; X_1,.......X_n) = \begin{bmatrix} g(W_1 X_1 + b_1)..... \; g(W_m X_m + b_m \\ . \qquad\qquad\qquad\qquad . \\ . \qquad\qquad\qquad\qquad . \\ g(W_1 X_n + b_1)......g(W_m X_n + b_m) \end{bmatrix} \tag{3.7.}$$

And can be stated as (Suresh ve ark., 2010):

$$\beta = \begin{bmatrix} \beta_1^T \\ . \\ . \\ \beta_m^T \end{bmatrix}_{mxm} \quad and \quad Y = \begin{bmatrix} Y_1^T \\ . \\ . \\ Y_n^T \end{bmatrix}_{nxm} \tag{3.8.}$$

where, $H$ is hidden layer output matrix. Training of a network in a traditional feedforward ANN means seeking a solution for least squares in linear equation of $H\beta = Y$ in ELM. ELM algorithm can be summarized in three stages as follows (Huang et al., 2010):

1. Stage: $W_i = (W_{i1}, W_{i2}, W_{i3}.....W_{in})$ input weights and hidden layer $b_i$ bias values are produced randomly
2. Stage: H hidden layer output is computed
3. Stage: $\hat{\beta}$ output weights are computed according to $\hat{\beta} = H^{+}Y$. Y is a decision feature.

### 3.7.3. Multiple linear regression

The investigation of various direct relapse (MLR) is a measurable strategy that inspect cause-impact connections amongst reliant and free factors. In MLR, the connection between info variable more than one (***x1, x2, . . .xn***) and a needy variable (y) is inspected. The relapse work that will be utilized here is characterized as take after:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + .... + \beta_n x_n \qquad (3.9)$$

Where it is acknowledged that every autonomous variable has a straight association with a needy variable (Civelekoglu et al., 2008).

The practical association amongst needy and free factor can be expressed with framework shape as underneath.

$$Y = X\beta + e \qquad (3.10)$$

Where Y is an output variable vector of size *nx1*; X is an input variable matrix of size *n x (p + 1); b* is a coefficient vector of size *(p + 1)x1* and e is an error vector of size *n x1*. According to Eq. (9), a variable multi-linear of p regression can be written as below;

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & . & X_{1p} \\ 1 & X_{21} & . & X_{2p} \\ . & . & . & . \\ 1 & X_{n1} & . & X_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} \qquad (3.11)$$

$\beta$ regression parameter coefficients in matrix can be showed as below;

$$\beta = (X^{'} X)^{-1} X^{'} Y \qquad (3.12)$$

Where $\beta$ regression coefficients are obtained through least square method (Apaydın et al., 1994).

### 3.7.4. Support vector machine

SVM is a pattern classification method based on the statistical learning theorem (Cortes and Vapnik, 1995). The goal in SVM is to find the most appropriate separator plane that classifies the data set as well as possible. That is, it is aimed to find the situation where the distance between two classes is the greatest. This objective is achieved by having the largest boundary between different samples after being transferred to the high dimension where the nonlinear sample space can be linearly separated (Vapnik, 1995). SVMs are divided into linear and non-linear SVMs. There is no need to use non-linear kernel functions in linear SVMs. In non-linear SVM, it is necessary to select an appropriate non-linear kernel function for probing.

#### 3.7.4.1. Linear support vector machines

Linear SVMs are examined as linearly separable and linearly non-separable SVMs. In the case of being able to be linearly separated, the samples $\{x_i, y_i\}, i = 1,2,...,N,$ representing these two classes should be separated by a separating plane directly to represent the $y_i \in \{-1,+1\},$ label values and the d-dimensional feature vector when defined as a set of N elements to be used for training. The purpose of the SVM is to separate the given data set with a lower plane according to the labels defined and leave all data points of the same class on the same side of the lower plane.

When it is assumed any x point on the separator plane, w separator plane normal and $|b|/\|w\|$ is the original perpendicular distance that the separating plane then the DVM algorithm tries to find optimal separator plane with below equation.

$$f(x) = w^T x + b = 0 \tag{3.13}$$

For this, the training set must have the following relation:

$$\begin{aligned} y_i &= +1 \text{ for, } w^T x_i + b \geq +1 \\ y_i &= -1 \text{ for, } w^T x_i + b \leq -1 \end{aligned} \tag{3.14}$$

When these inequalities are expressed together equation (3.15) is obtained.

$$y_i(w^T x_i + b) - 1 \geq 0 \qquad i=0,\ 1,\ \ldots,N \tag{3.15}$$

Here w and b values must be calculated to find the most suitable separating plane. Representative representation of the linear separable state is given in Figure 3.5(a).

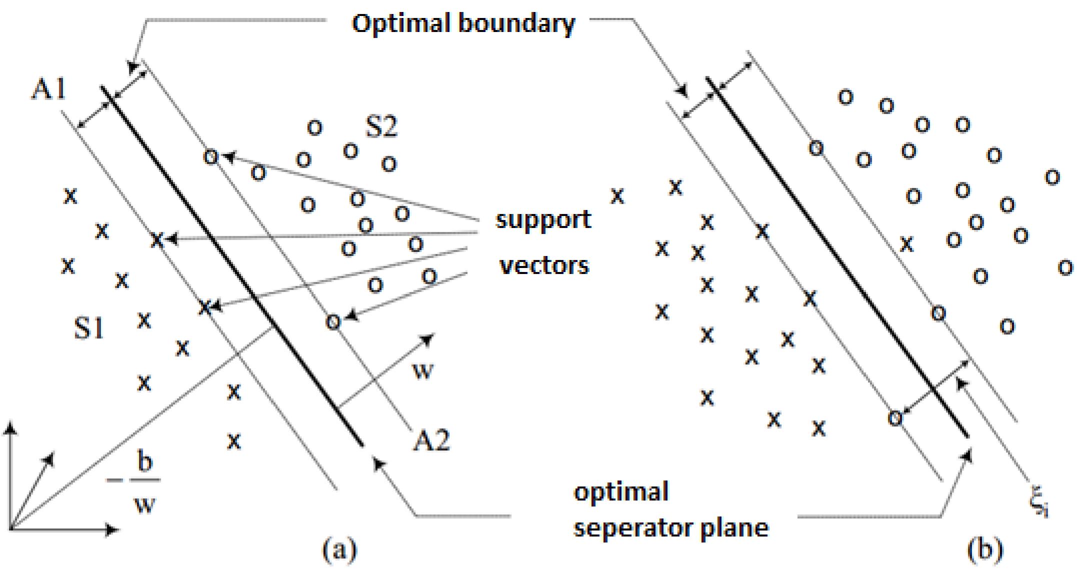


**Figure 3.5.** (a): Linear separation state, (b): Linear non-separable state

When it is assumed that the equation 3 forms the A1 separator plane that separates S1 class and equation 4 forms the A2 separator plane that separates class S2 in the same way; the perpendicular distance of the A1 to origin distance is $1-|b|/\|w\|$ and A2 separator plane to origin distance is $|-1-b|/\|w\|$. The optimal separation plane distances of these two separator planes are $|1|/\|w\|$. In other words, the distance between two sample clusters is $|2|/\|w\|$ because of the separator planes of A1 and A2 being parallel to each other. It should be noted here if there are no examples of training data between A1 and A2 separator planes. The greatest distance between these two separator planes can be found by reducing the $\|w\|$ value to the minimum.

The attempt is made by the SVM method to ensure that the distance between these two separator planes is the largest. The maximum boundary between these two planes is expressed by below equations.

$$\min \frac{1}{2}\|w\|^2 \tag{3.16}$$

$$y_i(w \cdot x_i + b) - 1 \geq 0 \tag{3.17}$$

Equation (3.16) represents the problem to solve, and equation (3.17) refers to the condition used during solution of the problem. Moreover, this expression is a second-order optimization problem. For solution of the problem Lagrange formulation can be applied. There are two reasons for the application of the Lagrange formulation. The calculation of the first Lagrange multipliers is easier. Secondly, it is better to generalize the problem for the nonlinear situation (Vapnik, 1995). The Lagrange formulation of the problem is;

$$L_p = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i y_i \left(w^T x_i + b\right) + \sum_{i=1}^{N} \alpha_i \tag{3.18}$$

The $\alpha_i \geq 0$ values in these formulas are called positive Lagrange multipliers. However, solving the formulation expressed in Eq. (3.18) is quite complex. To find the solution, equation (3.18) should be converted to dual probing using Karush-Kuhn-Tucker (KKT) conditions. For this problem, the solution depends on the requirements of KKT;

$$L_p = \sum_{i} \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j \tag{3.19}$$

The solution of equation (3.19) is carried out by the second order optimization problem under conditions $\alpha_i \geq 0$. Note that there is one Lagrange multiplier for each training instance. The vast majority of the Lagrange multipliers obtained in solution will be zero. The remaining $\alpha_i \geq 0$-valued $x_i$ examples are support vectors and are located on the A1 or A2 separator planes. Examples where the Lagrange multiplier is zero are the samples remaining on the back sides of the A1 or A2 separator planes. If the samples are not linearly separable, the positive weakness variables $\xi_i = 1,2,\ldots,N$ are used to solve the problem. The representation of the most appropriate separator plane representing this case is given in Figure 3.5 (b). If the conditions $\xi_i \geq 0$ in Eqs. (3.14) and (3.15) are to be redefined by the weakness variables should be as below.

$$y_i = +1 \text{ for, } w^T x_i + b \geq +1 - \xi_i \tag{3.20}$$

$$y_i = -1 \text{ for, } \mathrm{w}^T \mathrm{x}_i + \mathrm{b} \leq -1 + \xi_i \tag{3.21}$$

In the case of $\xi_i = 0$, the xi sample is correctly classified and the $\xi_i \geq 1$ is misclassified. If it can not be linearly separated, a C parameter is added to the system to prevent a solution to any possible situation within the training data. At the same time, this parameter also indicates the greatest value that the Lagrange multipliers can take. This allows Lagrange multipliers to remain in the $0 \leq \alpha_i \leq C$ range. The C configuration parameter is one of the parameters that the SVM should determine during the training phase. In this direction, the Lagrange formulation is rearranged as in equation (3.21):

$$L_p = \frac{1}{2}\|w\|^2 + C\sum_i \xi_i - \sum_i \alpha_i \left\{ y_i \left( w^T x_i + b \right) - 1 + \xi_i \right\} - \sum_i \mu_i \alpha_i \tag{3.21}$$

Where $\mu_i$ is a Lagrange parameter used to ensure that the weakness variables ($\xi_i$) remain positive. If the KKT conditions are applied for the solution of this Lagrange formulation,

$$L_d = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j \tag{3.22}$$

Obtained. Here, the $x_i$ values in the $0 < \alpha_i < C$ range, which correspond to the Lagrange multipliers, represent the support vectors. The way to find the solution in non-linear problems is to move to a space where the kernel functions and samples can first be separated in a higher dimensional and linear manner, and the solution is sought in this new space. This can be done by assuming $\Phi : \Re^d \mapsto H$ that L is the function φ which carries the d-dimensional feature space to an Euclidean space. Thus, the training algorithm of SVM will depend on the inner products. Here K represents the core function. Thus, the classifier can be determined by the decision function in Eq. (3.24.) for any unknown x instance after training.

$$\Phi(x_i) \cdot \Phi(x_j) = K(x_i, x_j) \tag{3.23}$$

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i \Phi(x_i) \cdot \Phi(x) + b = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b \tag{3.24}$$

In this function, N is the number of support vectors and xi is the support vectors. Some studies have been carried out in the literature on evaluating the classification performance of core functions in different fields. It is emphasized that high performance of classification can be achieved by using radial-based kernel functions in these studies (Schölkopf et al., 1997). Because radial basis kernel functions can exhibit both the sigmoid kernel function and the linear kernel function depending on the selected parameter ranges (Keerthi and Lin, 2003).

## 3.8. Performance Measures

### 3.8.1. Measures for classification problems

The precision, affectability and specificity criteria were utilized in order to demonstrate the execution of the associated systems. Precision rate recognizes the degree of correctly assembled cases to each one of the examples; affectability perceives the degree of accurately requested positive illustrations; specificity perceives absolutely described negative examples. These criteria will measure performance for classification problems for floor or building estimation.

These criteria are computed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \qquad (3.25)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \qquad (3.26)$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \qquad (3.27)$$

where *TP* (true positive) as well as *TN* (true negative) indicates true classifications and *FP* (false positive) as well as *FN* (false negative) indicates false classifications.

### 3.8.2. Measures for regression problems

In this survey, the estimation shows of the ELM, ANN and backslide models are taken a stab at using the going with authentic goof criteria: root mean square misstep (RMSE), mean inclination botch (MBE), and coefficient of affirmation (R2). The start quantifiable criteria are used to Show the presentations of the SVM, ELM, ANN and Regression exhibit in estimation of extension and longitude. Criteria are described by the going with conditions:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2} \qquad (3.28)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}[\hat{Y}_i - Y_i] \qquad (3.29)$$

$$R^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2 - \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \qquad (3.30)$$

where, Y is the genuine regard; is the mean of genuine regard; is the evaluated regard and n is the total number of perceptions. The MBE gives information general game plan execution of a model. A positive MBE addresses an overestimation while a negative MBE shows an underestimation. The RMSE gives information on the transient execution of a model. The estimation of RMSE is continually positive, tending to zero in the ideal case (Mother and Iqbal, 1983). The R2 can be used to pick the speedy connection between the consider and assessed regards (Bakirci, 2009). For perfect data showing up, RMSE and MBE should be more similar to zero, however estimation of R2 should approach to manage regulate 1 as almost as could sensibly be run of the mill.

## 4.RESULTS

The Experiments in the carried on in this study are composed of the categories, namely the part concerning classification and that of regression . The classification part is further composed of experiments the applied Extreme Learning Machine(ELM) for Regression, Support Vector Machines (SVM) and Artificial Neural Networks (ANN) to the estimation of floor and building using classification methods and Regression methods. The second part of the experiments Estimations of latitude and Longitude using classification methods and measures the performance using Root Mean Square Error (RMSE), MBE, $R^2$.

### 4.1. Experimental Implementation

To evaluate the performance of the proposed Machine Learning methods, experimental algorithms which include ELM, SVM and ANN and implemented and the testing of performance measures using RMSE, MBE, and $R^2$;

- The first step involves acquiring the UJIndoorLoc database which include both training and validation dataset.
- The data is normalized while features are extracted to perform experimental analysis using the stated machine leaning methods.

### 4.2. Dataset

UJIndoorLoc dataset contains a huge number of Wi-Fi localization fingerprints. The dataset contains a total of three buildings with 4 floors and 17 users participated in the experiment that generated the available data. Figure 4.1 shows the clustering of data before classification and is a 3D plot of Flow at the plane, Longitude and Latitude on the X and Y axis respectively. Figure 4.2 shows another 3D scatter plot of the three floors in building 2 before prediction.

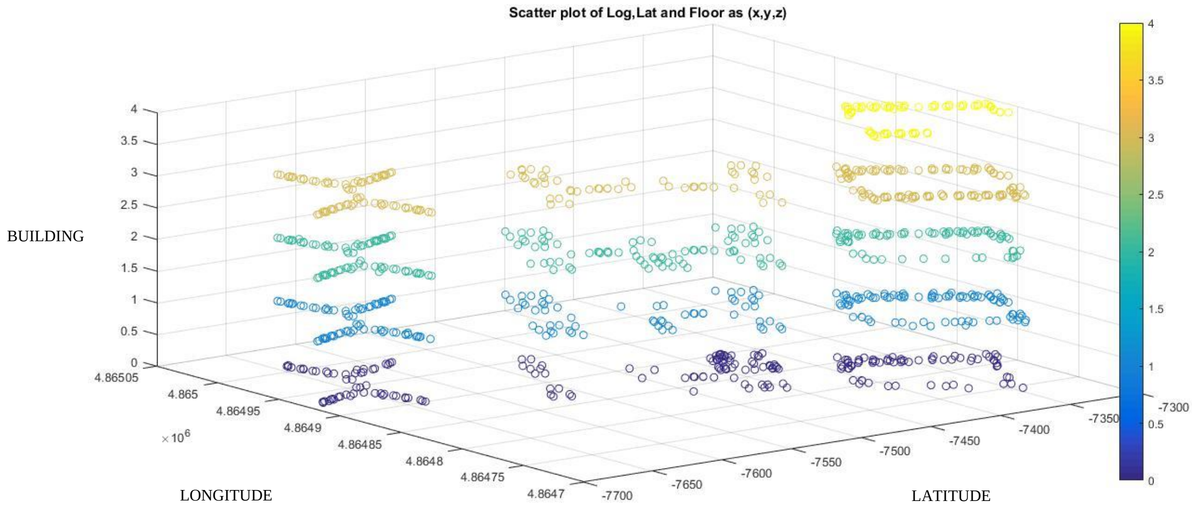


**Figure 4.1.** Scatter plot of the 3 buildings before prediction.

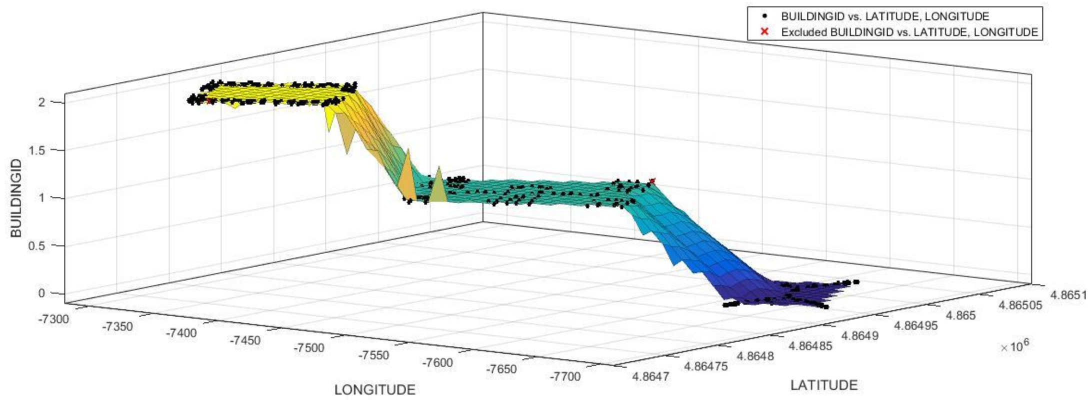


**Figure 4.2.** Scatter plot of floor of building 2 before prediction.

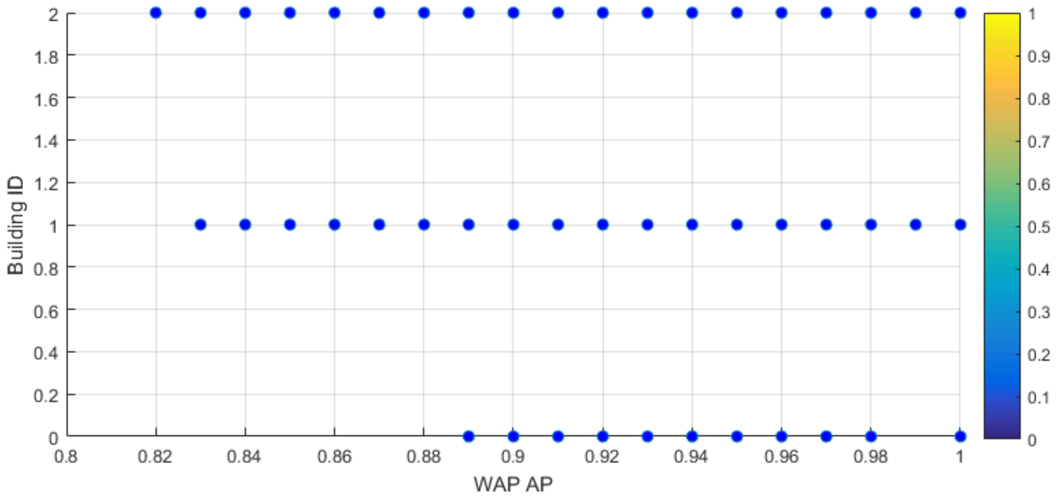


**Figure 4.3.** Result showing the positional distance of wifi on each floor.

According to figure 4.3 above, the buildings are plotted against the access point. The figure shows the linear location of the Wi-Fi in relation to the building. Also, it depicts the number of Wifi AP available in each building.

## 4.3. Classification

### 4.3.1 Estimation of floor using ELM

To estimate the floor, we made use of ELM and SVM as the estimation algorithms. Experiment conducted on the ELM algorithm show a very good performance of 99.99% testing accuracy as shown in the Table 4.1. It is evident from this table that the overall testing time was short, yet a very robust prediction accuracy was archived on the floor column of the dataset.

**Table 4.1.** ELM prediction accuracy result.

| TrainingTime | TrainingAccuracy | TestingTime | TestingAccuracy |
|---|---|---|---|
| 0.0781 sec | 99.88 | 0.0625 sec | 99.99 |

### 4.3.2. Estimation of floor using SVM

During the experimental stage, a classifier was trained to estimate the floor using Wi-Fi fingerprints. Support Vector Machine was used as the classifier and the recognition rate surpassed those of other algorithms used so far.

There is an accuracy of 91.9% using SVM and Euclidean distance matric. Recognition rate is as shown in table 4.2 below. Figure 4.4 shows the confusion matrix that shows the true positives and false negative accuracy rates of floor estimation using SVM.

**Table 4.2.** Floor estimation using the SVM classifier.

| **Classifier** | SVM |
|---|---|
| Accuracy | 91.9% |
| Prediction speed | ~91000 obs/sec |
| Distance Matric | Euclidean |


**Figure 4.4.** Confusion matrix showing SVM on floor estimation.

### 4.3.3. Classification of floor using ANN

The reliability R of the test was high using the artificial neural network, archiving up to the value of 99.77%. Figure 4.5 below describes the results in this experiment.

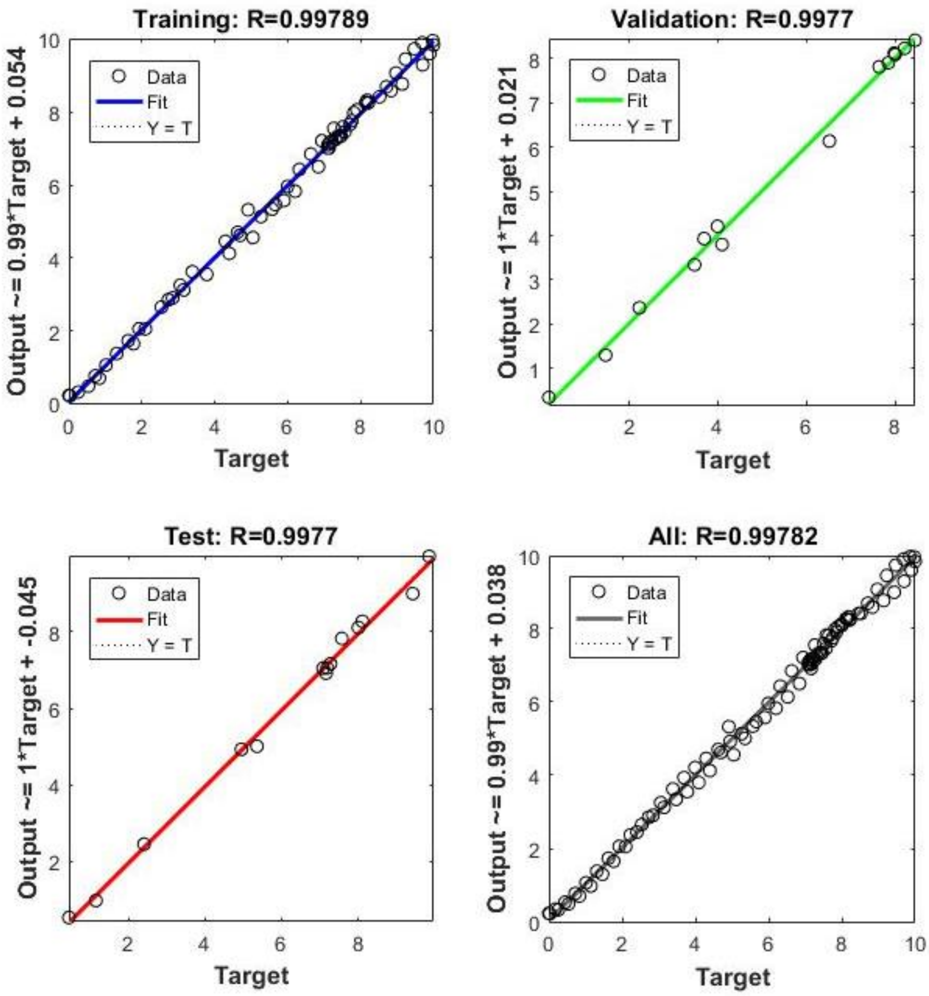


**Figure 4.5.** ANN classification floor.

### 4.4. Estimation of Building Using ELM

To estimate the building, we made use of ELM and SVM as the estimation algorithms. Experiment conducted on the ELM algorithm show a very good performance of 99.99% as shown in the Table 4.3. It is evident from this table that the overall testing time was short, yet a very robust prediction accuracy was archived on the floor column of the dataset.

**Table 4.3.** ELM prediction accuracy result.

| TrainingTime | TrainingAccuracy | TestingTime | Testing Accuracy |
|---|---|---|---|
| 0.1063 sec | 97.58 | 0.0313 sec | 99.99 |

### 4.5. Classificatioon of Building ANN

The reliability R of the classification on the building was high using the artificial neural network, archiving up to the value of 100%. Figure 4.6 below describes the results in this experiment.

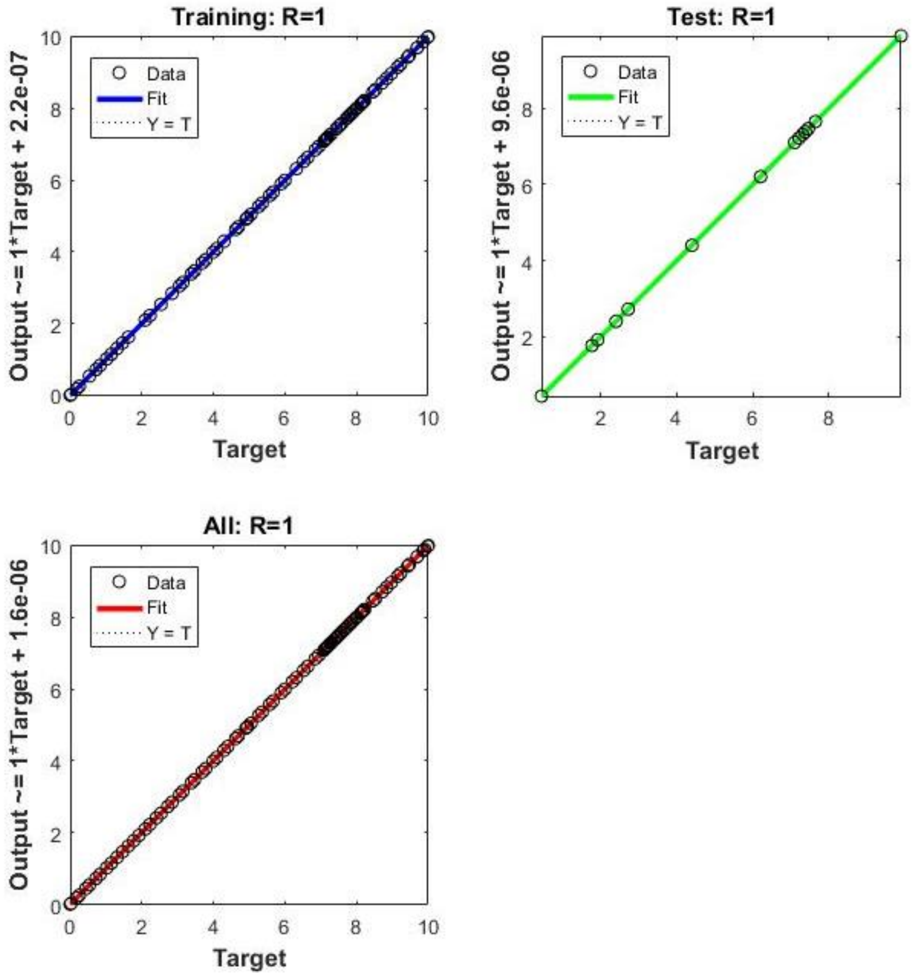


**Figure 4.6.** Classification of building using ANN.

### 4.6. Estimation of Building Using SVM

During the experimental stage, a classifier was trained to estimate the floor using Wi-Fi fingerprints. Support Vector Machine was used as the classifier and the recognition rate surpassed those of other algorithms used so far. There is an accuracy of 99.3% using SVM and Euclidean distance matric. Recognition rate is as shown in table 4.4 below.

Figure 4.7 shows the confusion matrix that shows the true positives and false negative accuracy rates of floor estimation using SVM. The number of neighbors were 100, distance metric is Euclidean, and distance weight is Equal.

**Table 4.4.** Building estimation using the SVM classifier.

| **Classifier** | SVM |
|---|---|
| Accuracy | 99.3% |
| Prediction speed | ~11000 obs/sec |
| Distance Matric | Euclidean |

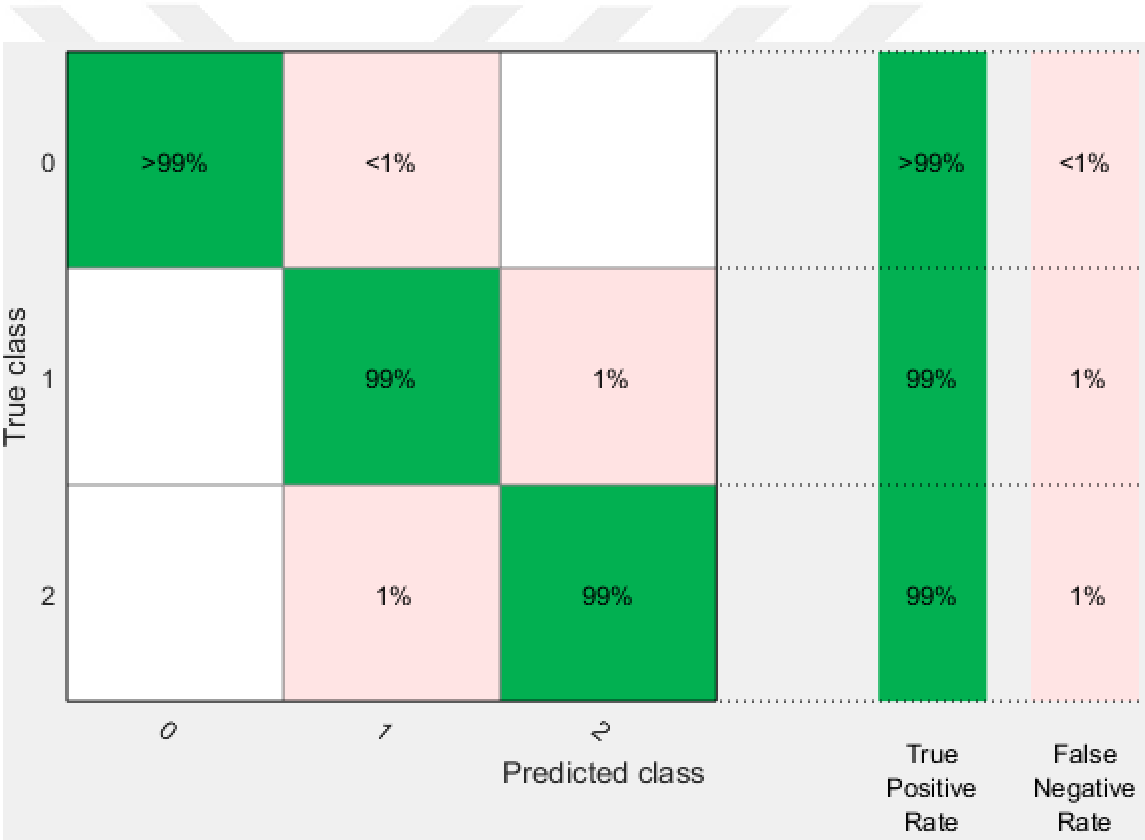


**Figure 4.7.** Confusion matrix showing SVM on building estimation.

## 4.7. Regression Analysis

Regression analysis been a process in statistics, used to estimate how variables are related to each other and how many techniques used in analyzing and modeling to discover the relationship between dependent and non-dependent variables otherwise known as predictors (Seber, et al, 2012).

The experiments carried out in this section of the thesis, analyzes the longitude and latitude of the Wi-Fi localization fingerprint based on the normalized data from UJIndoorLoc dataset.

### 4.7.1. Estmation of latitude

Here, Linear model Polynomial with formula as represented in equation 4.1:

$$f(x) = p1*x + p2 \tag{4.1}$$

In the above function, x is normalized by mean 48.65 and standard deviation std 0.0006693. We archived a confidence bound of 95% confidence bounds on these coefficients p1 and p2. Where p1 = -1.062 (-1.07, -1.053) and p2 = -74.64 (-74.65, -74.63). Table 4.5 summarizes the result of this section's experiment on latitude estimation using regression analysis.

**Table 4.5.** Estimation of latitude.

| | |
|---|---|
| Mean | 48.65 |
| Std | 0.0006693 |
| confidence bounds | 95% |
| R-square | 0.7401 |
| RMSE | 0.6291 |
| Classifier | SVM |

Locally weighted smoothing linear regression:

f(x,y) = lowess (linear) smoothing regression computed from p where x is normalized by mean -7470 and std 116.1 and where y is normalized by mean 4.865e+06 and std 67.39 Coefficients: R-square: 0.1225, adjusted R-square: 0.1108, RMSE: 1.161.

### 4.7.2. Application of linear regression on the floor position

Here, General Gaussian model with formula as represented in equation 4.2 is used to find the floor position:

$$f(x) = a1*exp(-((x-b1)/c1)^2) \tag{4.2}$$

In the above function, x is normalized by mean 48.65 and standard deviation std 0.0006693. We archived a confidence bound of 95% confidence bounds on these coefficients a1, b2 and c1. Where a1, b2 and c3 are summarized in table 4.5.

**Table 4.6.** Estimation of Floor, regression analysis.

| Reference point (RP) | |
|---|---|
| | Matrix form (x,y) |
| 1 | a1 =1.95 (1.923, 1.977) |
| 2 | b2 =91.75 (91.56, 91.95) |
| 3 | c3= 8.597 (8.153, 9.042) |
| | |
| Confidence bounds | 95% |
| SSE | 5.371e+04 |
| R-square | 0.02898 |
| Adjusted R-square | 0.02888 |
| RMSE | 1.641 |

### 4.7.3. Regression analysis using polynomial function

Here, Polynomial function model with formula as represented in equation 4.3 is used to find the user position.

$$f(x,y) = p00 + p10*x + p01*y \tag{4.3}$$

In the above function, is a polynomial function plotted on userID against the AP, in order to determine user . We archived a confidence bound of 95% confidence bounds on these coefficients p00, p10 and p01. Where a1, b2 and c3 are summarized in table 4.5.

**Table 4.7.** Estimation of users on each building and floor, regression analysis.

| Reference point (RP) | |
|---|---|
| | Matrix form (x,y) |
| 1 | p00 = 8.39 (7.993, 8.787) |
| 2 | p10 = -0.08166 (-0.08587, -0.07745) |
| 3 | p01 = 0.0534 (0.05129, 0.05552) |
| | |
| Confidence bounds | 95% |
| SSE | 1.143e+04 |
| R-square | 0.1742 |
| Adjusted R-square | 0.1741 |
| RMSE | 0.7571 |

## 4.8. Estmation of Latitude Using ANN

The reliability R of the test was high using the artificial neural network, archiving up to the value of 99%. Figure 4.8 below describes the results in this experiment.

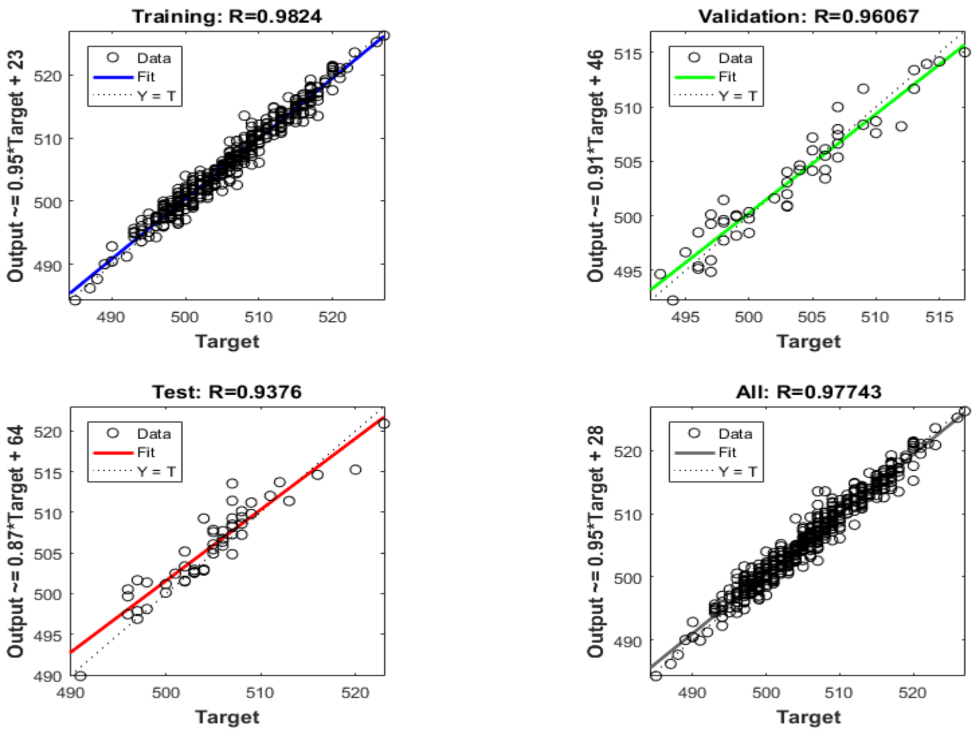


**Figure 4.8.** ANN result of the latitude.

## 4.9. Estmation of Longitude Using ELM

Table 4.8. It is evident from this table that the overall testing time was short, yet a very robust prediction accuracy was archived on the floor column of the dataset where 83.73% accuracy was archived on the testing result.

**Table 4.8.** ELM prediction accuracy result for longitude.

| TrainingTime | TrainingError | TestingTime | Testing Error |
|---|---|---|---|
| 1.6875 secs | 0.7882 | 0.0469 sec | 0.8073 |

## 4.10. Estmation of Longitude Using SVM

During the experimental stage, a classifier was trained to estimate the floor using Wi-Fi fingerprints. Support Vector Machine was used as the classifier and the recognition rate surpassed those of other algorithms used so far. There is an accuracy of 92.5% using SVM and Euclidean distance matric. Recognition rate is as shown in table 4.9. below.

**Table 4.9.** Building estimation using the SVM classifier.

| **Classifier** | SVM |
|---|---|
| Accuracy | 92.5% |
| Prediction speed | ~02100 obs/sec |
| Distance Matric | Euclidean |

### 4.11. Estmation of Longitude Using ANN

Estimation of The reliability R of the longitude was high using the artificial neural network, archiving up to the value of 91%. Figure 4.9. below describes the results in this experiment.

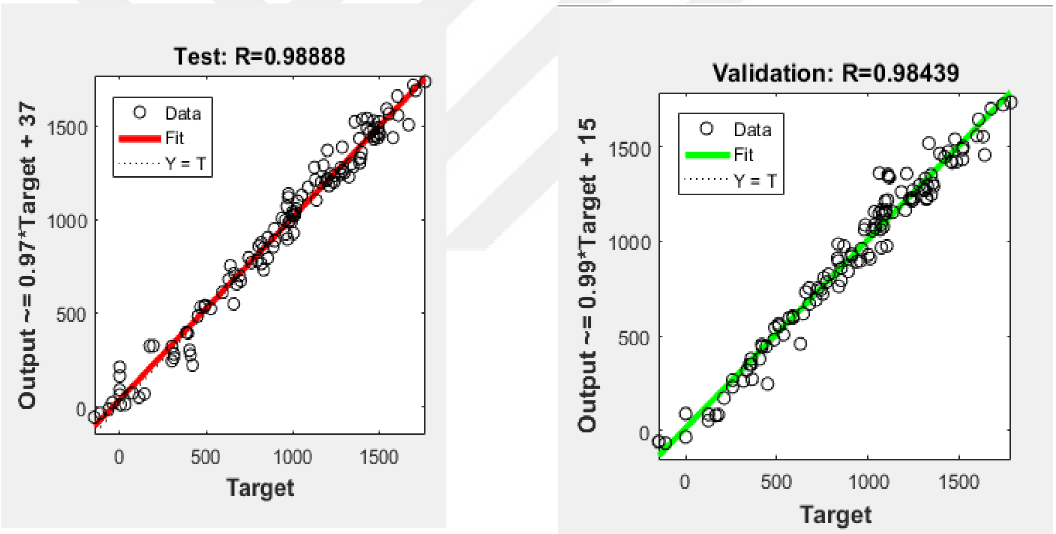


**Figure 4.9.** ANN result of the longitude.

# 5. CONCLUSIONS AND FUTURE WORK

## 5. Conclusions

Wi-Fi indoor po positioning relies on the wireless technology of Wi-Fi systems in order to acquire the indoor location information, which is of great significance to the development of indoor positioning applications. This recommendation is generally based on how to upgrade the current standard arranging counts and further proposes an improved Wi-Fi indoor arranging estimation by weighted mix. The proposed estimation relies on upon the ordinary region fingerprinting count. By using the Wi-Fi hail botch dealing with, better fingerprints in the midst of the separated getting methodology could be picked up". In the wake of improving the customary Euclidean division arranging and the joint probability arranging, a more correct zone result is expert. This hypothesis furthermore taken a gander at the changed machine learning figurings used.

In this study, the performances of ELM, SVM and regression approach was been evaluate and it has greatly performed in this location with low percentage of error. There are several task can be extended in the future works. Firstly, the classifier can be improve by consider the user is moving, this is because many Wi-Fi fingerprints are clustered on the same position. This is because the RSS samples of a stationary user can be collected for as many time as needed to improve the accuracy, but for moving user, a very few RSS samples of location can be collected. “Exploration of the big variations in historical RSS information because to enhance the system estimation, RSS pattern plays important role in determination location is highly suggested. Irregular RSS patterns will decrease the accuracy of the estimation location. Furthermore, the accuracy of estimation location can be increases as increase the measurements of RSS sampling thus provide a better system performance.

However, appeared differently in relation to other channel-related fingerprints, which require submitted devices and high transmission limit, this new one of a kind finger impression is more insightful since it can be capably gotten from channel estimation comes to fruition that exist in most present day remote specific contraptions. We change the remarkable finger impression into the logarithmic scale to ensure that the parts of the exceptional finger impression vector contribute sensibly to the zone estimation. By multiplications, we have shown that the mix of logarithmic-scale Wi-Fi novel finger impression with RSS demonstrates better execution advantage when contemplated than other ordinary finger impression based techniques, and moreover the arrangement which joins neural framework and isolated components. The centrality of progress in precision is affirmed under different information transmission conditions. Propagation comes to fruition have also exhibited that, our proposed plan is not quite recently solid to persistent channel assortments realized by unpredictable positions and presentations of human bodies, however in the meantime is more beneficial in utilizing gear structure and get ready effort, stood out from various arrangements proposed in the written work.

**5.2. Future Work**

In the proposed computation, other than arranging precision, decreasing work costs is a key issue for examiners as well. It is similarly the future heading for this proposed work. To diminish work costs, the future work can be revolved around the going with headings. Immediately, the best way to deal with build up the database of fingerprints actually, In addition, once the database of fingerprints has been set up, the issue of keeping up it without human obstacle should be clarified. Thirdly, the counts of part limit with all extraordinary check vectors in the planning database display genuine overhead in the limitation strategy, especially when more than one customer ought to be discovered at the same time. We thusly require a shrewd database looking for technique to diminish the time spent on looking. Likewise, our work in this paper focuses on the errand of discovering static customers. Before long, the customers are moving now and again. Making use of the steady assortment of the channel based information for flexible customer taking after will be a trying errand.

## 6. REFERENCES


Addesso, P et al. 2013. “An Indoor Localization System Within an IMS Service Infrastructure.” *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*: 321:321--321:325. http://doi.acm.org/10.1145/2536853.2536926.

Arackaparambil, Chrisil, Sergey Bratus, Anna Shubina, and David Kotz. 2010. “On the Reliability of Wireless Fingerprinting Using Clock Skews.” : 169–74.

Bacak, Ahmet, and Hasari Çelebi. 2014. “Al I Nan Sinyal Gücü RF Parmakizi Tabanl I Binaiçi Konumland I Rma Sistemlerinde Pratik Faktörler Practical Considerations For RSS RF Fingerprinting Based Indoor Localization Systems.” (Siu): 497–500.

Bahl, Paramvir, and Venkata N Padmanabhan. 2000. “RADAR: An In-Building RF-Based User Location and Tracking System.” *Proc. IEEE INFOCOM 2000. The 19th annual conference on Computer Communications* 2: 775–84.

Bekkelien, Anja. 2012. “Bluetooth Indoor Positioning.” (March).

Berkvens, Rafael, Maarten Weyn, and Herbert Peremans. 2015. “Mean Mutual Information of Probabilistic Wi-Fi Localization.” *Indoor Positioning and Indoor Navigation (IPIN), 2015 International Conference on* (October): 4.

Box, P O, and Klong Luang. 2010. “Indoor Positioning System Using Artificial Neural Network Hamid Mehmood , Nitin K . Tripathi and Taravudh Tipdecho Department of Remote Sensing and Geographic Information Systems , School of Engineering and Technology , Asian Institute of Technology ,.” 6(10): 1219–25.

Bratus, Sergey, Cory Cornelius, David Kotz, and Daniel Peebles. 2008. “Active Behavioral Fingerprinting of Wireless Devices.” : 56–61.

Calderoni, L., Ferrara, M., Franco, A., & Maio, D. (2015). Indoor localization in a hospital environment using random forest classifiers. Expert Systems with Applications, 42, 125–134.

“Chapter 1 , Wireless System Architecture : How Wireless Works.” 2008.

Chen, Bo, Vivek Yenamandra, and Kannan Srinivasan. 2015. “Tracking Keystrokes Using Wireless Signals.” *Mobisys*: 31–44.

Chen, Mike Y Sohn, Timothy Chmelev, Dmitri Haehnel, Dirk Hightower, Jeffrey Hughes, Jeff LaMarca, Anthony Potter, Fred Smith, Ian Varshavsky, Alex/ Practical metropolitan-scale positioning for gsm phones.Ubiquitous Computing, 225-242, 2006.

Cheng, J., Yang, L., Li, Y., & Zhang, W. (2014). Seamless outdoor/indoor navigation with wifi/gps aided low cost inertial navigation system. Physical Communication, 13, Part A, 31-43. Indoor Navigation and Tracking.

Civelekoglu, G., et al. "Prediction of bromate formation using multi-linear regression and artificial neural networks." *Ozone: Science and Engineering* 29.5 (2007): 353-362.

Deng, Z. A., Wu, D., Zhou, Y., & Na, Z. (2016). Enhanced Local Fisher Discriminant Analysis for Indoor Positioning in Wireless Local Area Network.Future Internet, 8(2), 8.

Deng, Zhi-an, Yubin Xu, and Liang Chen. 2013. "Localized Local Fisher Discriminant Analysis for Indoor Positioning in Wireless Local Area Network." (2012): 4795–99.

Evans, Joseph B., Weichao Wang, and Benjamin J. Ewy. 2006. "Wireless Networking Security: Open Issues in Trust, Management, Interoperation and Measurement." *International Journal of Security and Networks* 1(1/2): 84. http://www.inderscience.com/link.php?id=10825.

Ferris, Brian Fox, Dieter Lawrence, Neil D. WiFi- SLAM Using Gaussian Process Latent Variable Models. IJCAI, 7: 2480–2485, 2007.

"Fingerprint-Based Technique for Indoor Positioning System Via Machine Learning and Convex Optimization." 2016.

"GPS Use in U . S . Critical Infrastructure and Emergency Communications Presented to the United States Technical Training Institute ( USTTI )."

Guo, Y. , Chen, Y. , & Liu, J. (2014). Indoor location estimation based on local magnetic field via hybrid learning. In F. Sun, K.-A. Toh, M. G. Romay, & K. Mao (Eds.), Extreme learning machines 2013: Algorithms and applications . In Adaptation, learning, and op- timization, Vol. 16 (pp. 189–207). Springer International Publishing.

Gupta, Priya, and S S Sutar. 2014. "Study of Various Location Tracking Techniques for Centralized Location , Monitoring & Control System." 4(3): 27–30.

Huang, G.-B., Zhu, Q.-Y. ve Siew, C.-K., 2004. Extreme learning machine: A new learning scheme of feed forwardneural networks, IEEE International Joint

Conference on Neural Networks, 2, 985–990.

Huang, G.B., Zhu, Q.Y. ve Siew, C.K., 2006. Extreme learning machine: Theory and applications, Neurocomputing, 70, 489-501.

Husen, Mohd Nizam, and Sukhan Lee. 2016. "High Performance Indoor Location Wi-Fi Fingerprinting Using Invariant Received Signal Strength." *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication - IMCOM '16*: 1–6. http://dl.acm.org/citation.cfm?doid=2857546.2857589.

Idris, L. I., Hashim, N. M. Z., JA'afar, A. S., Salleh, A., Mohamad, N. R., & Aziz, K. A. A. (2006). INDOOR LOCATION ESTIMATION UTILIZING WI-FI SIGNAL A. (2006).

Ijaz, F. , Yang, H. K. , Ahmad, A. , & Lee, C. (2013). Indoor positioning: A review of indoor ultrasonic positioning systems. In Proceedings of the advanced communication technology (pp. 1146–1150).

Ijeh, Anthony C, Allan J Brimicombe, David S Preston, and Chris O Imafidon. 2009. "Security Measures in Wired and Wireless Networks." : 113–21.

Jiang, P., Zhang, Y., Fu, W., Liu, H., & Su, X. (2015). Indoor mobile localization based on wi-fi fingerprint's important access point. International Journal of Distributed Sensor Networks, 2015, 45.

Joaquín Torres-Sospedra, Raúl Montoliu, Adolfo Martínez-Usó, Tomar J. Arnau, Joan P. Avariento, Mauri Benedito-Bordonau, Joaquín Huerta UJIIndoorLoc: A New Multibuilding and Multi-floor Database for WLAN Fingerprint-based Indoor Localization Problems In Proceedings of the Fifth International Conference on Indoor Positioning and Indoor Navigation, 2014.

Karem, C., Taha, B.M.J.Q., Stuart, H.,Hosni, G.M. ve Hugo, G.,2008. Comparison of ice-affected streamflow estimates computed using artificial neural networks and multiple regression techniques, Journal of Hydrology ,349,383– 396.

Kaya, Yılmaz, and Lokman Kayci. "Application of artificial neural network for automatic detection of butterfly species using color and texture features." *The visual computer* 30.1 (2014): 71-79.

Kaya, Yılmaz, et al. "Evaluation of texture features for automatic detecting butterfly species using extreme learning machine." *Journal of Experimental & Theoretical*

*Artificial Intelligence* 26.2 (2014): 267-281.

Keerthi, S.S. ve Lin, C.-J., 2003. Asymptotic behaviors of support vector machines with gaussian kernel, Neural Computation, 15(7), 1667–1689.

Kimura, Yosuke. 2012. "Development of Wall Amusement with Infrared Radars." : 367–70.

Kuo, Y. S. , Pannuto, P. , Hsiao, K. J. , & Dutta, P. (2014). Luxapose: Indoor positioning with mobile phones and visible light. In Proceedings of the 20th annual international conference on mobile computing and networking MobiCom '14 (pp. 447–458). New York, NY, USA: ACM .

Lan, K. C., & Shih, W. Y. (2013). On calibrating the sensor errors of a pdr-based indoor localization system. Sensors, 13, 4781–4810.

Lee, Kyogu. 2014. "Integration of a Precise Indoor Position Tracking Algorithm with an HMD-Based Virtual Reality System." : 23–26.

Lemiux, Nathan, and Hanan Lutfiyya. 2009. "WHLocator : Hybrid Indoor Positioning System." : 55–63.

Li, Dan, Wang Le, and Shiqi Wu. 2012. "Indoor Positioning System Using Ultrasonics." : 440–44.

Li, H. (2014). Low-cost 3d bluetooth indoor positioning with least square. Wireless Personal Communications, 1–14.

Liang, N. Y., Huang, G.-B., Saratchandran, P. ve Sundararajan, N., 2006. A fast and accurate online sequential learning algorithm for feedforward networks, IEEE Transactions on Neural Networks, 17, 1411–1423.

Liu, Shu, Yingxin Jiang, and Aaron Striegel. 2014. "Face-to-Face Proximity Estimation Using Bluetooth On Smartphones." 13(4): 811–23.

Living, Ambient-assisted, Óscar Belmonte-fernández, Adrian Puertas-cabedo, and Joaquín Torres-sospedra. 2017. "An Indoor Positioning System Based on Wearables." : 1–22.

Ma, Rui, Qiang Guo, Changzhen Hu, and Jingfeng Xue. 2015. "An Improved WiFi Indoor Positioning Algorithm by Weighted Fusion." : 21824–43.

Marques, Nelson Meneses, Filipe Moreira, Adriano. Combining similarity functions and majority rules for multi-building, multi-floor, WiFi positioning. IEEE Xplore, 2012.

Martí, J. V. , Sales, J. , Marín, R. , & Jiménez-Ruiz, E. (2012). Localization of mobile sensors and actuators for intervention in low-visibility conditions: The zigbee fingerprint- ing approach. International Journal of Distributed Sensor Networks, 2012 , 1–10.

Mascharka, David, and Eric Manley. “Machine Learning for Indoor Localization Using Mobile Phone-Based Sensors.”

Mathisen, Andreas et al. “A Comparative Analysis of Indoor WiFi Positioning at a Large Building Complex.”

Mirowski, P., Whiting, P., Steck, H., Palaniappan, R., MacDonald, M., Hartmann, D., & Ho, T. K. (2012). Probability kernel regression for WiFi localisation. Journal of Location Based Services, 6(2), 81-100.

Montaser, A., & Moselhi, O. (2014). RFID indoor location identification for construction projects. Automation in Construction, 39, 167–179.

Piontek, Hubert, and Æ Matthias Seyffer. 2007. “Improving the Accuracy of Ultrasound-Based Localisation Systems.” : 439–49.

“RADAR_in_building_RF.pdf.”

Rashid, Tahmid, Shadman Sakib Chowdhury, and Mehbas Fairuz Nawal. 2016. “Composite Indoor Localization and Positioning System Based on Wi-Fi Repeater Technology.” : 2130–33.

Saeid, Mirzaei Azandaryani (Florida Atlantic University). 2013. “INDOOR LOCALIZATION USING WI-FI FINGERPRINTING.” Florida Atlantic University Boca Raton, Florida.

Sanzgiri, Kimaya et al. 2002. “A Secure Routing Protocol for Ad Hoc Networks.” *Icnp*: 78–89. http://computer.org/proceedings/icnp/1856/18560078abs.htm.

Schölkopf, B., Sung, K., Burges, C., Girosi, F., Niyogi, P., Poggio, T. ve Vapnik, V., 1997. Comparing support vector machines with Gaussian kernels to radial basis function classifiers, IEEE Trans. on Signal Processing, 45(11), 2758–2765.

Seber, G.A. and Lee, A.J., 2012. *Linear regression analysis* (Vol. 936). John Wiley & Sons.

Shi, Ke et al. 2015. “Support Vector Regression Based Indoor Location in IEEE 802 . 11 Environments.” 2015.

States, United. 2004. “Sanjay Sarma, Oatsystems and Mit 50.” (October).

Suresh S., Saraswathi, S. ve Sundararajan, N., 2010. Performance enhancement of extreme learning machine for multi-category sparse data classification problems, Eng. Appl. of Artificial Intelligence, 23, 1149–1157.

Thiem, Lasse, Björn Riemer, Marcus Witzke, and Thomas Luckenbach. 2008. “Poster Abstract : RFID-Based Localization in Heterogeneous Mesh Networks.” : 415–16.

Torres-Sospedra, Joaquín et al. 2015. “Enhancing Integrated Indoor/outdoor Mobility in a Smart Campus.” *International Journal of Geographical Information Science* 8816(July): 1–14. http://www.tandfonline.com/doi/full/10.1080/13658816.2015.1049541.

Torres-Sospedra, J., Montoliu, R., Trilles, S., Belmonte, Ó. & Huerta, J. (2015). Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems. Expert Systems with Applications, 42(23), 9263-9278.

Tsui, A. W., Chuang, Y. H., & Chu, H. H. (2009). Unsupervised learning for solving RSS hardware variance problem in WiFi localization. Mobile Networks and Applications, 14(5), 677-691.

Ugave, Viney Anand. 2014. “SMART INDOOR LOCALIZATION USING MACHINE LEARNING TECHNIQUES Submitted.” *Department of Electrical and Computer Engineering In*.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory, Springer-Verlag, New York.

Weng, Dongdong, Yetao Huang, Yue Liu, and Yongtian Wang. 2009. “Study on an Indoor Tracking System with Infrared Projected Markers for Large-Area Applications.” 1(212): 239–46.

Woo, Sunkyu Jeong, Seongsu Mok, Esmond Xia, Linyuan Choi, Changsu Pyeon, Muwook Heo ,Joon. Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR. Automation in Construction, 20: 3-13, 2011.

Zandbergen, Paul A. 2012. “Comparison of WiFi Positioning on Two Mobile Devices.” 6(1): 35–50.

Zhou, Junyi Shi, Jing. RFID localization algorithms and applications: a review. Journal of Intelligent Manufacturing, 20: 695–707, 2009.

## CURRICULUM VITAE

Yadgar Ahmed HASSAN was born in 1987 in Halabja, Iraq. He completed his primary and secondary education in Halabja, Iraq. He studied Department of Communication Engineering at Technical College of Engineering (TCE) of Sulaimani Polytechnic University(SPU) in Sulaimani/Iraq. He graduated Bachelor in 2012©. He started in Turkey to study M.Sc. at Siirt University institute of science Electrical and Electronics Engineering Department in 2015©. He can speak Kurdish, English, Arabic and Turkish.