Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 1, April 2023, pp. 911-921
ISSN: 2302-9285, DOI: 10.11591/eei.v12i2.4200

Bluetooth beacons based indoor positioning in a shopping malls using machine learning

Kamel Maaloul[1], Brahim Lejdel[2], Eliseo Clementini[3] , Nedioui Med Abdelhamid[2]
[1]LABTHOP Laboratory, Department of Informatic, Faculty of Exact Sciences, University of El-Oued, El-Oued, Algeria
[2]Department of Informatic, Faculty of Exact Sciences, University of El-Oued, El-Oued, Algeria
[3] Department of Industrial and Information Engineering and Economics, University of L'Aquila, L'Aquila, Italy

Article Info

Article history:

Received Jun 4, 2022
Revised Aug 14, 2022
Accepted Oct 30, 2022

Keywords:

Bluetooth beacons
Extra-trees classifier
Indoor localization
Machine learning
Smartphone sensors

ABSTRACT

The adoption of Bluetooth beacon technology demonstrates a broad interest in indoor positioning technology because of its low cost and ease of use. Bluetooth beacons usually have an accuracy of fewer than 4 meters. The use of machine learning (ML) leads to results with greater accuracy compared to using traditional filtering methods. In this paper, we provide indoor localization based on Bluetooth beacons using several different ML techniques. We used ML algorithms to locate customers' devices in shopping malls. The extra-trees classifier and k-neighbors classifier found the device with greater than 90% accuracy. Other algorithms were able to determine the location with less accuracy. The results also showed that Bluetooth technology is a valid solution to find the data used to analyze the spatial-temporal behavior of individuals.

Corresponding Author:

Kamel Maaloul
LABTHOP Laboratory, Department of Informatic, Faculty of Exact Sciences, University of El-Oued
El-Oued PB 789, El-Oued, Algeria
Email: maaloul-kamel@univ-eloued.dz

1. INTRODUCTION

Indoor positioning methods are in great demand nowadays for accurate positioning [1]. The signals from the global positioning system (GPS) cannot accurately determine indoor locations. Because walls greatly impede signal strength, making it impossible for them to pass through tall buildings and move within structures. Indoor location can be determined through the development of indoor positioning technologies, such as radio frequency identification (RFID), wireless accuracy (WiFi), and Bluetooth [2]. The indoor positioning system helps determine information about the movement patterns of customers. It also offers advanced customer service, such as the most visited places for shoppers in the mall and the reorganization of public facilities [3].

Market visitors can be dealt with by a device that works on Bluetooth and WiFi instead of a guide. These technologies are the most widely used because of their low cost and high accuracy, and they are used in all smartphones. We use Bluetooth-compatible devices because of their generally small size, low battery consumption, and low cost. This is done by sending a globally unique identifier and then picking it up by a compatible application or operating system. Received signal strength indication (RSSI) measurements represent the relative quality of a received signal on a device. After accounting for potential antenna and cable level losses, the RSSI shows the power level received. The stronger the signal, the higher the RSSI value. The number that is nearer to zero when measured in negative numbers typically indicates a better signal. The position calculation is based on RSSI values [4].

The effectiveness of the indoor location system requires important things, including permanent and accurate identification of the location of the user, rapid identification of the site; and the ability to adapt to a changing and difficult environment [5]. The development of artificial intelligence algorithms has increased the handling of machine learning (ML) techniques. It also led to the growth of data to enhance the quality and effectiveness of services provided to customers in malls [6]. ML algorithms can be used to classify data into different categories or predict regressions with a continuous variable [7].

iBeacon devices communicate with smart devices to send wireless signals. Each iBeacon device has a unique identifier that represents a store. It automatically saves customer data, including a unique identifier and time data. This information can track the customer's path in the shopping center [8]. This is to understand customers' pain points. They can shift the focus to their needs, allowing them to create more effective and satisfying experiences. To solve the problem of finding a user's location using Bluetooth RSSI values, ML techniques are used to enhance the accuracy of the patriot. This is because RSSI values are unstable even within the same distance owing to the influence of elements in the surrounding environment such as weather, humidity, physical barriers, and interference from other signals.

In our research, ML models quantitatively use the available datasets as test datasets. Each site is classified based on its proximity to the access point. The ML model also predicts the location in the test data set. The distance between the device's initial location and the particular device is the basis for studies in other papers. In our study, we label the locations and then use several ML-based methods to train the models to classify and discover the locations [9]. We also analyzed and compared the performance of 6 individual predictors using ML algorithms in indoor localization features. Then they evaluated model performance by measures of accuracy, precision, and recall in ML.

Currently, a different set of works that depend on ML approaches for indoor localization of websites have been presented. Therefore, it determines indoor locations using collaborative positioning techniques, which rely on the exchange of information between different users and/or devices to improve the overall situation of the system. There are different algorithms used to achieve positionings, such as fingerprints, multilateralism, or triangulation. Other systems exist based on wireless signals, and optical or magnetic solutions [10]. The researchers used the fingerprint as a method of positioning by analyzing different algorithms and techniques. The wide variety of mistakes that result in the heterogeneity of devices and the complexity of environmental variables present one of the biggest obstacles in indoor positioning [11].

Firdaus *et al*. [12] decreased k-nearest neighbor (KNN) search's computational tim. So when the value of k in KNN is greater, the computation time increases, especially when using the Cityblock and Minkowski space functions. According to Handojo *et al*. [13], visitors can find out where they are in the museum and how long they have spent there. Internal GPS using BLE signals is used for this mapping. The exhibit room has BLE beacons placed at specific locations. The program that the museum visitor has loaded on their phone picks up the signals from the beacons. Research by Duong *et al*. [14], use the triangulation method and Kalman filter provided by the program to determine the locations of visitors. Also, use the BLE beacon to increase the accuracy of the internal GPS, which uses triangulation.

To precisely measure the distance, the suggested method uses the RSSI range (higher than 70 dB, which is equivalent to a distance of fewer than 3 meters). Communicating the predicted position of the tripods to dependable circuits improves positioning accuracy. Additionally, the four beacons' combined power is employed for a more precise location. By utilizing the suggested algorithm and basing their research on two well-known AP devices, Mosleh *et al*. [15] looked into localization. Several receiving points have been deployed around the floor of the entire room for testing and measuring the RSS signal.

To address the shortcomings of the RF-wireless communication standards, Lokanatha *et al*. [16] devised a digital HBC transceiver (TR) hardware architecture that adheres to the IEEE 802.15.6 standard. A frequency-selective digital transmission system is used in the design. Through the use of various field programmable gate array (FPGA) families, the design resources are examined. Muharam *et al*. [17] focused on the length of the training period. The fixed target parameter (FTP) and shifting target parameter techniques are used throughout the training phase (MTP). MTP was 5 seconds faster than FTP in terms of the time needed to obtain RSSI data from each reference node. The beacon-based campus management system is intended to be built using the layered architecture by Dong *et al*. [18]. The suggested architecture makes use of Bluetooth low energy 4.0 beacon technology, which enables data sharing through Bluetooth at very low power consumption—using just one coin cell battery can last for years.

The goal of this study is to provide an accurate and compared assessment of indoor localization utilizing a variety of ML models in a novel and distinctive manner, building on prior research. After collecting the data in a specific location using smartphones. The results are analyzed and compared to determine the best performing and most accurate algorithm based on the distance and RSSI values, although they are different. This paper tries to introduce what is different than other research studies. We propose a positioning algorithm for an indoor positioning system using Bluetooth and ML. We compared the performance in terms of the mean of error and evaluated model performance by measures of accuracy, precision, and recall in ML.

2. THE PROPOSED METHOD

The major objective of the suggested model for indoor positioning is to gather information about the parameters involved in indoor positioning, such as distance and position, using this information to train the model. The chart below shows the proposed model for Bluetooth beacons-based indoor positioning in Shopping Malls using several ML algorithms. The suggested indoor positioning system is made up of three main parts. The beacon RSSI values that mobile devices have received are first collected by a data collecting module. The filter algorithm corrects the data after that, and a data processing module uses the filtered findings to apply a positioning method [19]. The data management module, which is the third main component, stores and manages the data processing results in a database. In Figure 1, the suggested indoor positioning system is displayed.

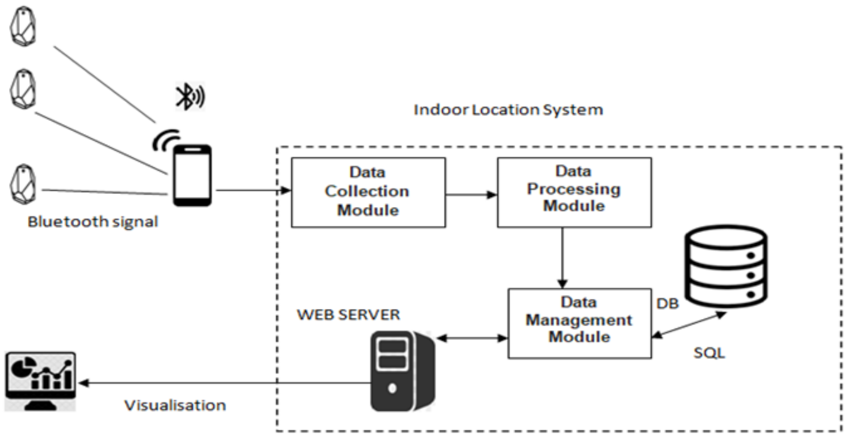

Figure 1. Architecture of the proposed indoor positioning system

2.1. Data collection module

From the user's mobile device, the data collection module gathers a variety of data. A mobile application on the user's device uses Bluetooth connectivity to gather data from a beacon and deliver it to the indoor positioning system's data collection module. The message authentication code (MAC) value of the device and the beacon's identifier are transmitted to the data acquisition module via the mobile application during communication with the mobile device.

2.2. Data processing module

The RSSI data obtained from the data acquisition module is corrected by the data processing module using a filtering technique to limit the error range. This module determines the distance between each beacon and the mobile device using the RSSI data of the beacons that were gathered by corresponding with those devices in the collection module. It then uses those distances to determine the position of the user in the interior space.

2.3. Data management module

After the data processing module has examined the information gathered by the acquisition module, the data management module determines the location of a user's mobile device and handles various data needed by the indoor positioning system. The following system stages serve as the foundation for ML algorithms: the first stage involves preprocessing all of the data in a set to categorize it using ML classifiers; the second stage involves classifying the data, and the third stage involves using ML algorithms and determining results [20]. As a result, ML approaches are used in a variety of sectors, including marketing, medicine, and so on. Classification and regression problems can both benefit from ML algorithms. Indoor location classification techniques are the best choice for predicting discrete data.

Several KNN, RFC, extra trees classifiers (ETC), SVM, gradient boosting classifiers (GBC), and decision trees (DT) algorithms have been chosen. This is due to the fact that they are employed with continuous variables that are assumed to be feature-independent. Different input properties can predict one or more output values. Because of the large amount of data that we will train and with the presence of notes more than features, low bias/high variance techniques such as KNN, DT, or SVM were also used. These algorithms, which accept a longer time to train than other times to train, can also achieve high accuracy. A brief description of the categorization algorithms used in this work is provided:

2.3.1. K-neighbors classifier

The principle of K-neighbors classification (KNC) is to locate a predetermined number, i.e., the k of training samples that are closest in distance to a fresh sample that needs to be classified. The new sample's label will be determined by its neighbors. User-defined constants control how many neighbors must be determined in KNN classifiers [15]. Based on point density, radius-based neighbor learning algorithms have a variable number of neighbors; all samples within a defined radius have a variable number of neighbors. An important ML algorithm is the KNN algorithm. It has excelled in numerous applications for regression and classification [21].

a. Decision tree

DT are a common ML approach that creates a tree-like structure to represent decisions. They're created using measurements like gene impurity and information to create a top-down framework. Both classification and regression problems are modeled using DT [22]. The DT is simple, but it over-splits into traits and learns with the training data critically. To avoid this, they are usually trimmed to stop them from becoming any larger.

b. Random forest classifier (RFC)

Random forests (RF) are a type of ensemble learning method for classification, regression, and other problems that work by building a large number of DT during training [23]. Regression and classification are only two of the many issues that can be solved using the potent ML method known as RF. Because RF model uses an ensemble technique, it is composed of numerous tiny DT, or estimators, each of which generates its own predictions. Scikit-learn uses averaging for higher accuracy and over-fitting management rather than asking each tree classifier to vote for a label. The averaging strategy does not alleviate bias in the classifier's output, but it can provide lower variance than a DT classifier [24].

c. Extra trees classifier

ETC and FC are two ways of grouping that are extremely similar. However, there is a difference. The ETC uses the entire original sample, whereas the RFC uses bootstrap replicas, which means it subsamples the input data with replacement [25]. There is an optional parameter in the Extra Trees sklearn implementation that allows users to bootstrap replicas, but it defaults to using the whole input sample. Because bootstrapping diversifies the data, this may increase variance. Another distinction is the use of cut points for splitting nodes. The optimum split is determined by the RFC, whereas the ETC chooses it at random. The two methods evaluate the best among the subset of features after choosing the split points [26].

d. Support vector machine classifier (SVMC)

SVMs are supervised learning models and their corresponding learning algorithms are used in ML for regression and classification analyses. A SVM training method, a non-probabilistic binary linear classifier, creates a model that categorizes fresh data measurements into one of two groups [27].

e. Gradient boosting classifier

A group of ML techniques known as GBC combine numerous weak learning models to create a powerful predictive model. DT are frequently used for gradient enhancement. In order to carefully identify the ideal arrangement of trees, the gradient boosting method sequentially produces basic models from a weighted version of the training data. The goal of each basis model addition phase is to fix the errors created by the preceding base models. Consequently, the gradient boosting approach has the potential to deliver forecasts that are more precise [28].

3. METHOD

The method for indoor positioning was studied and applied to collect RSSI values and location coordinates (x, y) from stationary signals. The selected models were tested and compared based on their performance using performance measures. In this study, we address the indoor positioning problem of smartphones using Bluetooth signal localization as a classification problem.

3.1. Dataset

In our suggested approach, we used Bluetooth beacons to predict the indoor position using smartphones. Data collection from multiple access points (AP) on the building's floor was the initial phase. 3 Kontakt Bluetooth beacons are mounted in a 2.74 m wide×4.38 m long (width×length) area of the building.

The 3 beacons are transmitting at a transmit power of -12 dbm. A Sony Xperia E3 smartphone with Bluetooth turned on is used as a receiver to record the data. Recordings are done in several positions in the building of an interval of 30-60 seconds in the same position. The dataset generated after data collection contains several signal strengths at different locations. Comma-separated values (CSV) format has been used to transform the data. Initially, CSV files had characteristics like (distance a, distance b, distance c, position X, position Y, date, and time). Figure 2 shows the location of the Bluetooth and the structure of the building.

Figure 2. Location of the Bluetooth and the structure of the building

3.2. Data preprocessing

Preparing (cleaning and arranging) raw data in order to make it suitable for creating and training ML models is known as data preprocessing in ML. The dataset is filtered to remove noise using the running average, and if there is no RSSI value in between, the filter inserts the lowest possible RSSI value into that vacancy. The data set is also checked for null values, which are replaced by the lowest possible value or the above value in the data row [29]. Table 1 shows an example of the dataset. It is a sample of how observations were collected.

Table 1. Example of the dataset

Distance A	Distance B	Distance C	Position X	Position Y	Date	Time
0.877462	0.768608	1.457214	122	180	Feb 09 2017	12:20:22
1.201608	1.03122821	1.893498	122	180	Feb 09 2017	12:20:23
1.614344	1.098873	2.112560	122	180	Feb 09 2017	12:20:24
1.513634	1.135640	2.296293	122	180	Feb 09 2017	12:20:28
1.517499	1.148356	2.388172	122	180	Feb 09 2017	12:20:29
1.489097	1.163176	2.498214	122	180	Feb 09 2017	12:20:31
1.533095	1.174496	0.000000	122	180	Feb 09 2017	12:20:32

Fields:
- The distance in meters between beacon A and the device is calculated by using the RSSI of this Bluetooth beacon. (Distance A)
- The distance in meters between beacon B and the device is calculated by using the RSSI of this Bluetooth beacon. (Distance B)
- The distance in meters between beacon C and the device is calculated by using the RSSI of this Bluetooth beacon. (Distance C)
- X coordinate in centimeters rounded to the nearest centimeter measured using a meas-uring tape with +/-1 cm accuracy. (Position X)
- Y coordinate in centimeters rounded to the nearest centimeter measured using a meas-uring tape with +/-1 cm accuracy. (Position Y)

Figure 3 depicts the suggested categorization system's workflow. In most cases, one selects a portion of the pertinent population for which values of the target attribute are known or, if necessary, creates the data. The choice of a ML algorithm that will be utilized to suit the intended target number happens concurrently with this process. The majority of the effort entails creating, locating, and cleaning the data to make sure it is accurate, consistent. The second step is to determine how to map the system's properties—the model's input—in a form that is appropriate for the algorithm of choice. This entails converting the raw data into specific properties that will serve as algorithmic inputs. Once this procedure is complete, the model is trained by maximizing performance, which is often gauged using a cost function of some kind. Typically, this includes modifying the hyperparameters that regulate the model's training procedure, internal structure, and characteristics. The data are divided up into different sets. To optimize the hyperparameters, a validation dataset that is distinct from the test and training sets should be employed.

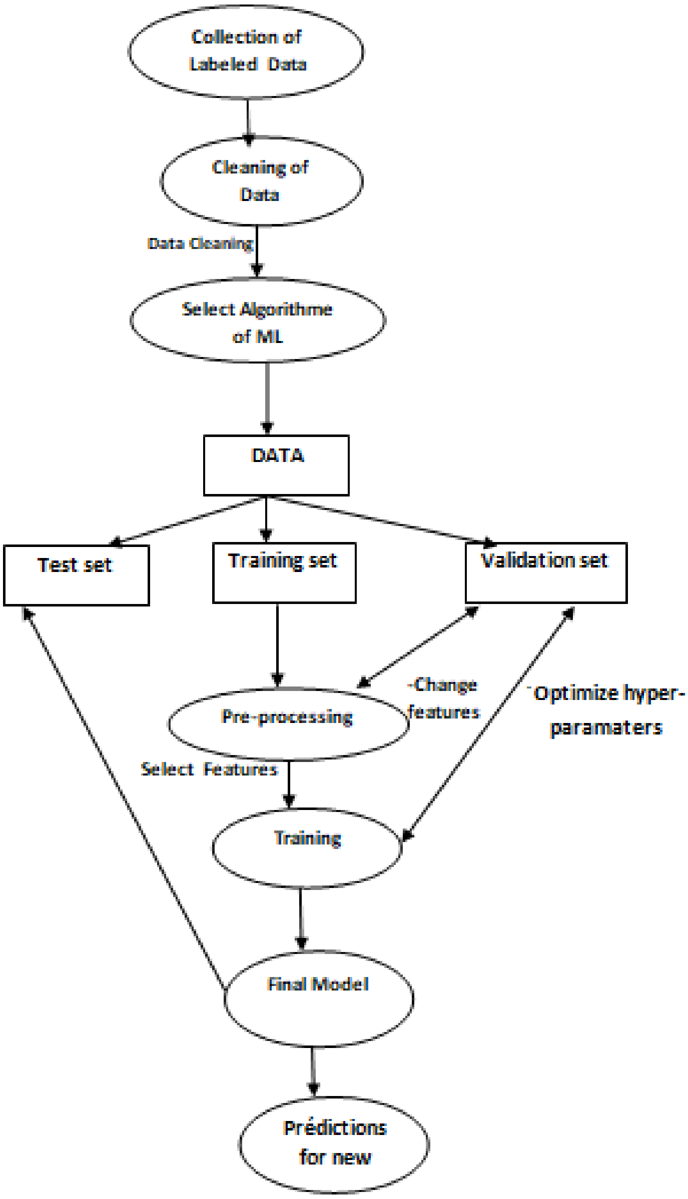

Figure 3. Workflow of the proposed classification system

Even graphs are produced when conventional graphs are extrapolated to larger dimensional data sets. Exploring correlations between multidimensional data sets with this is helpful. Principal component analysis is a linear dimensionality reduction technique that we have used to extract data from a high-dimensional space by projecting it into a low-dimensional subspace (PCA). Given that the data includes a lot of features and ML algorithm learning is quite slow, you can use this to reduce the training and testing times for ML algorithms. Figure 4 displays the main component analysis using Scikit-Learn.

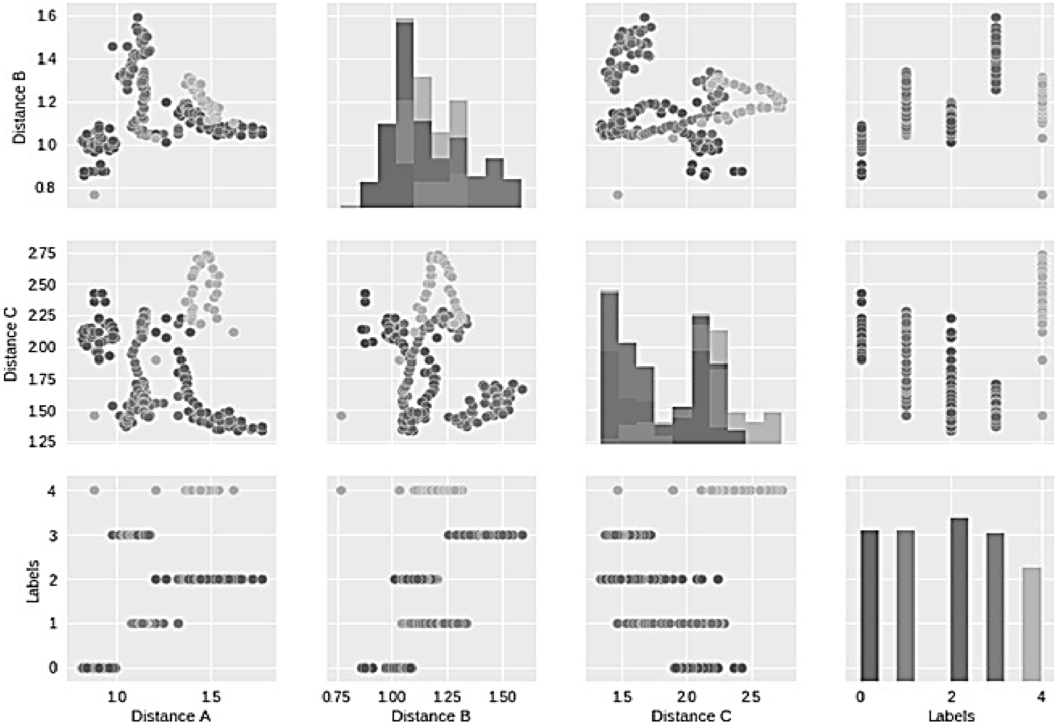

Figure 4. Principal component analysis using Scikit-Learn

A training set comprising 80% of the dataset and a testing set comprising 20% were then created. Although a dataset can be split into a variety of ratios, the experiment used the fairest splitting suggested for a small dataset [30]. Figure 5 displays data visualization in 3D. In order to create a better and more accurate data representation, it was created using a three-dimensional chart.

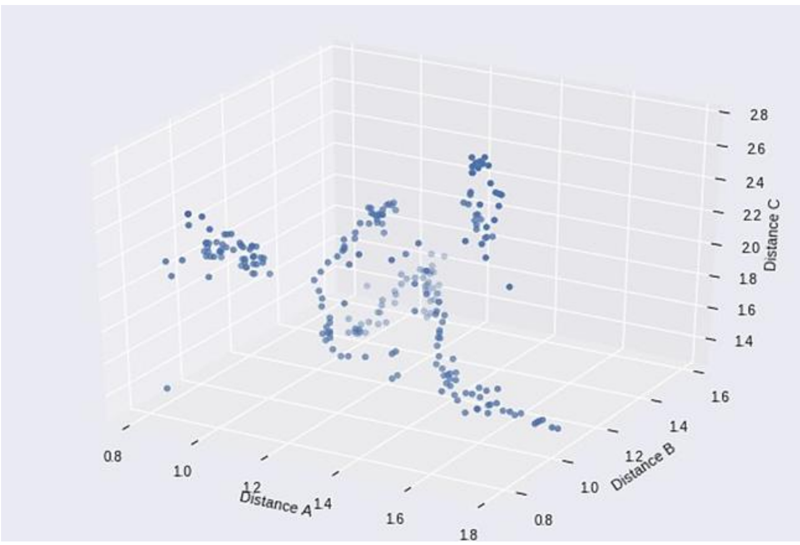

Figure 5. 3D plotting of data visualization

4. RESULTS AND DISCUSSION

4.1. Results

Although distance and RSSI values are varied, the experiment's results were examined and contrasted to identify the top-performing algorithm. Therefore, processing them separately and coordinating them is the ideal approach. Distance, (X, Y) coordinates, and RSSI values are examples of independent variables. An Intel i3-9100 with 16 GB of RAM and an LGA 1151 card served as the test PC. Six models built on KNN, RFC, ETC, SVM, GBC, and DT were trained. We delivered testing data sets for prediction to each of the six ML models.

A confusion matrix is a technique for summarizing and describing the performance of classification algorithms on a set of test data for which the true values are known [31]. The accuracy of classification is sometimes misleading due to the difference in the number of observations in each category or the multiplicity of categories in the data set [32]. In our work, we evaluated model performance by measures of accuracy, precision, and recall in ML. In this experiment, we used six algorithms, and the goal is to better predict where people are:

- The easiest performance metric to understand is accuracy, which is just the proportion of properly predicted observations to all observations. It is a great metric but only when we have similar data sets where the values of false positives and false negatives are nearly the same [33]. Therefore, we have to look at other parameters to evaluate the performance of the proposed model.
- Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. High accuracy is related to low false positive rate ie how many correct places are actually?
- Recall is the ratio of correctly predicted positive observations to the all observations in actual class.

The detection rate of each model based on the testing dataset and performance comparison of models can be seen in Table 2. Through the histogramme, ETC is an efficient algorithm capable of detecting the location of the device with a detection rate of 95.63%. The prediction performance when using different Models is in the Figure 6. The ETC algorithm made more accurate predictions in most shopping malls than other algorithms.

Table 2. Performance comparison of models

Model name	Accuracy	Precision	Recall (%)
KNC	92.24	91.31	91.20
RFC	88.66	88.02	87.90
ETC	95.63	94.85	94.80
SVC	82.63	81.55	81.55
GBC	89.96	89.60	89.51
DT	87.96	87.91	87.88

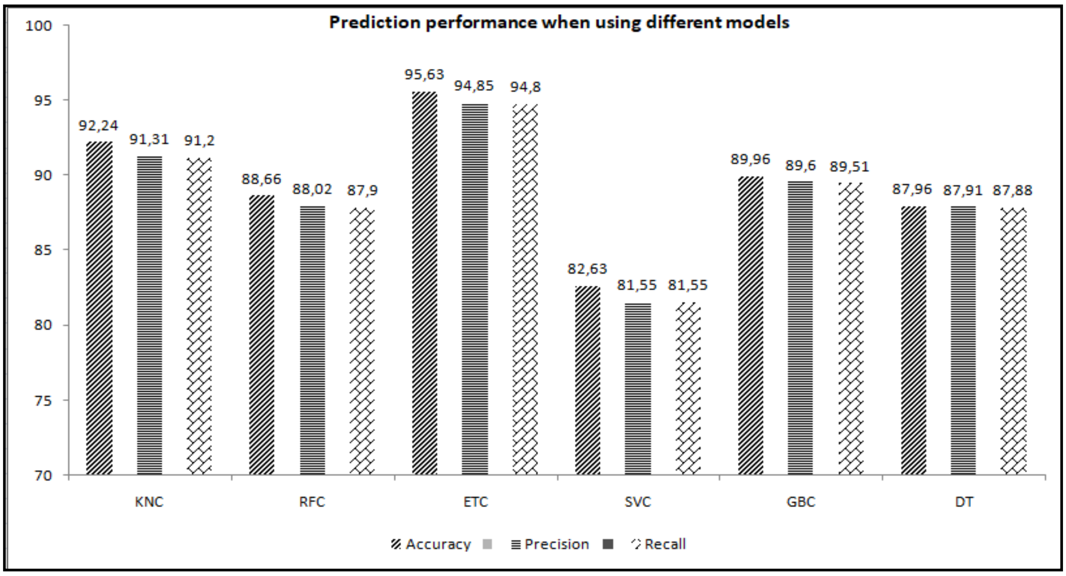

Figure 6. Prediction performance when using different models

We notice that the six algorithms in particular were able to accurately identify the majority of the locations based on the experiments done and the results attained. Microsites that have been sampled can also be found using ETC. The training of the data preparation models has been highly successful.

ETC outperforms KNN, GBC, RFC, and DT because it takes in more information based on the distance from the internet connection stage. Note that SVM performs worst because it needs more test samples for accurate localization. Using the SVM classifier is one of the weakest classification results because it has superior performance capability with small training sample sizes. The comparison experiment found that it is able to predict the location of the device with the best level of accuracy using detection rate as well as classification matrix. As shown in Table 2 and Figure 6.

4.2. Discussion

It's necessary to develop a localization system that can achieve high levels of precision in building-scale real-world environments while leveraging low-cost and widely available technologies like smartphones and Bluetooth devices. We created a dataset with values from Bluetooth RSSI and a training model to predict the distance from a specific beacon [34]. Numerous factors contribute to the ETC algorithm's exceptional performance. However, the fact that Bluetooth Beacons do not substantially rely on any particular set of capabilities is the main factor in the performance improvement of Bluetooth Beacons. Excellent results are obtained when the data collection process includes a large number of features. Excellent results are obtained when the considered methods need further improvements to reduce the required resources. Given the high computational requirements of ML algorithms, the experimental results showed the necessity of a sophisticated path loss model for RSSI-based distance estimation. In general, proximity accuracy can be greatly improved by filtering technical elements. The experimental results showed a significant improvement over the homogeneous results, achieving a convergence error range up to a few distances from the receiver. BLE signals are a promising solution because they are inexpensive, easy to deploy, and have low power requirements [35].

Wireless devices that operate in the radio spectrum now have the potential to interfere with BLE beacons' signals. It will be necessary to investigate performance over a longer period of time and to include additional users and places in the research [36]. By minimizing signals that could contribute to an incorrect prediction, we were able to anticipate a user's location with greater accuracy. Low signal strength may also help to extend the battery life of the beacons. Testing our model in a bigger area with a more complicated environment and more Bluetooth signals is one of our work's limitations. Our system could be expanded by including an additional stage as one potential means of enhancing scalability.

5. CONCLUSION

In this work, we analyzed and compared the performances of 6 individual predictors using ML algorithms in indoor localization features to characterize places on earth. We have validated the performance of the system using Bluetooth beacons on smartphones. We find that the six algorithms were able to accurately identify the majority of the locations. The ETC model based on Bluetooth beacons gave us the most accurate places based on the results of the experiments. Although we do good processing of the data, the ML algorithms require a lot of improvement because of the complex computations.

In the future, we will improve the data collection validation procedure and combine this work with an indoor tracking system to locate individuals with excellent accuracy. This gives us better results for the algorithm to keep track of anything needed indoors. We also plan to expand the range of the experiment to include multiple floors and define AP to increase accuracy. Our plan also includes working on a method for determining the access point, which can increase the accuracy of indoor localization.

REFERENCES

[1] H. Huang and L. L. Liu, "Analysis of the location of Nanning Large-Scale Mall based on BP neural network," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-3/W10, pp. 975–979, Feb. 2020, doi: 10.5194/isprs-archives-xlii-3-w10-975-2020.

[2] T. Pei, Y. Liu, H. Shu, Y. Ou, M. Wang, and L. Xu, "What Influences Customer Flows in Shopping Malls: Perspective from Indoor Positioning Data," *ISPRS International Journal of Geo-Information*, vol. 9, no. 11, p. 629, Oct. 2020, doi: 10.3390/ijgi9110629.

[3] J. Hrad *et al.*, "Indoor Positioning System Based on Fuzzy Logic and WLAN Infrastructure," *Sensors*, vol. 20, no. 16, p. 4490, Aug. 2020, doi: 10.3390/s20164490.

[4] A. Mussina and S. Aubakirov, "Improving Indoor Positioning via Machine Learning," in *Proceedings of the 8th International Conference on Data Science, Technology and Applications*, 2019, doi: 10.5220/0007916601900195.

[5] K. Konstantinos and T. Orphanoudakis, "Bluetooth Beacon Based Accurate Indoor Positioning Using Machine Learning," in *2019 4th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, Sep. 2019, doi: 10.1109/seeda-cecnsm.2019.8908304.

[6] Z. Ullah, F. Al-Turjman, L. Mostarda, and R. Gagliardi, "Applications of Artificial Intelligence and Machine learning in smart cities," *Computer Communications*, vol. 154, pp. 313–323, Mar. 2020, doi: 10.1016/j.comcom.2020.02.069.

[7] A. H. Salamah, M. Tamazin, M. A. Sharkas, and M. Khedr, “An enhanced WiFi indoor localization system based on machine learning,” in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct. 2016, doi: 10.1109/ipin.2016.7743586.
[8] J. Kang, J. Seo, and Y. Won, “Ephemeral ID Beacon-Based Improved Indoor Positioning System,” *Symmetry*, vol. 10, no. 11, p. 622, Nov. 2018, doi: 10.3390/sym10110622.
[9] T. Dornan, G. O’Sullivan, N. O’Riain, E. Stueeken, and R. Goodhue, “The application of machine learning methods to aggregate geochemistry predicts quarry source location: An example from Ireland,” *Computers &amp$\mathsemicolon$ Geosciences*, vol. 140, p. 104495, Jul. 2020, doi: 10.1016/j.cageo.2020.104495.
[10] F. Khelifi, A. Bradai, A. Benslimane, P. Rawat, and M. Atri, “A Survey of Localization Systems in Internet of Things,” *Mobile Networks and Applications*, vol. 24, no. 3, pp. 761–785, Aug. 2018, doi: 10.1007/s11036-018-1090-3.
[11] F. Liu *et al.*, “Survey on WiFi-based indoor positioning techniques,” *IET Communications*, vol. 14, no. 9, pp. 1372–1383, Jun. 2020, doi: 10.1049/iet-com.2019.1059.
[12] F. Firdaus, N. A. Ahmad, and S. Sahibuddin, “Fingerprint indoor positioning based on user orientations and minimum computation time,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 17, no. 4, p. 1740, Aug. 2019, doi: 10.12928/telkomnika.v17i4.12774.
[13] A. Handojo, T. Octavia, R. Lim, and J. K. Anggita, “Indoor positioning system using BLE beacon to improve knowledge about museum visitors,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, p. 792, Apr. 2020, doi: 10.12928/telkomnika.v18i2.14886.
[14] N.-S. Duong and T.-M. D. Thi, “Smartphone indoor positioning based on enhanced BLE beacon multi-lateration,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 1, p. 51, Feb. 2021, doi: 10.12928/telkomnika.v19i1.16275.
[15] M. F. Mosleh, R. M. Zaal, and E. I. Abbas, “Hybrid localization algorithm based on received signal strength and angle-of-arrival for indoor location estimation,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 2, p. 454, Apr. 2021, doi: 10.12928/telkomnika.v19i2.18324.
[16] S. B. Lokanatha and S. B. B. Prashanth, “Design and performance analysis of human body communication digital transceiver for wireless body area network applications,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, p. 2206, Jun. 2022, doi: 10.11591/ijece.v12i3.pp2206-2213.
[17] A. Muharam, A. Wahab, and M. Alaydrus, “Improved fingerprinting performance in indoor positioning by reducing duration of the training phase process,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 1, p. 236, Oct. 2021, doi: 10.11591/ijeecs.v24.i1.pp236-244.
[18] H. J. Dong, R. Abdulla, S. K. Selvaperumal, S. Duraikannan, R. Lakshmanan, and M. K. Abbas, “Interactive on smart classroom system using beacon technology,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, p. 4250, Oct. 2019, doi: 10.11591/ijece.v9i5.pp4250-4257.
[19] N. S. C. J, N. H. A. Wahab, N. Sunar, S. H. S. Ariffin, K. Y. Wong, and Y. Aun, “Indoor Positioning System: A Review,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 6, 2022, doi: 10.14569/ijacsa.2022.0130659.
[20] N. Nasir *et al.*, “Water quality classification using machine learning algorithms,” *Journal of Water Process Engineering*, vol. 48, p. 102920, Aug. 2022, doi: 10.1016/j.jwpe.2022.102920.
[21] M. Luo, J. Zheng, W. Sun, and X. Zhang, “WiFi-based Indoor Localization Using Clustering and Fusion Fingerprint,” in *2021 40th Chinese Control Conference (CCC)*, Jul. 2021, doi: 10.23919/ccc52363.2021.9549410.
[22] U. Singh, M. Rizwan, M. Alaraj, and I. Alsaidan, “A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step towards Smart Grid Environments,” *Energies*, vol. 14, no. 16, p. 5196, Aug. 2021, doi: 10.3390/en14165196.
[23] D. P. Mohandoss, Y. Shi, and K. Suo, “Outlier Prediction Using Random Forest Classifier,” in *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2021, doi: 10.1109/ccwc51732.2021.9376077.
[24] J. Lee, C. Park, and H. Roh, “Revisiting Adaptive Frequency Hopping Map Prediction in Bluetooth with Machine Learning Classifiers,” *Energies*, vol. 14, no. 4, p. 928, Feb. 2021, doi: 10.3390/en14040928.
[25] E. K. Ampomah, Z. Qin, and G. Nyame, “Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement,” *Information*, vol. 11, no. 6, p. 332, Jun. 2020, doi: 10.3390/info11060332.
[26] M. W. Ahmad, J. Reynolds, and Y. Rezgui, “Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees,” *Journal of Cleaner Production*, vol. 203, pp. 810–821, Dec. 2018, doi: 10.1016/j.jclepro.2018.08.207.
[27] Z. Zhao, J. Carrera, J. Niklaus, and T. Braun, “Machine Learning-Based Real-Time Indoor Landmark Localization,” in *Lecture Notes in Computer Science*, Springer International Publishing, 2018, pp. 95–106, doi: 10.1007/978-3-030-02931-9_8.
[28] S. Deb, A. K. Goswami, R. L. Chetri, and R. Roy, “Prediction of Plug-in Electric Vehicle’s State-of-Charge using Gradient Boosting Method and Random Forest Method,” in *2020 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Dec. 2020, doi: 10.1109/pedes49360.2020.9379906.
[29] I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters, and A. Ng, “Cybersecurity data science: an overview from machine learning perspective,” *Journal of Big Data*, vol. 7, no. 1, Jul. 2020, doi: 10.1186/s40537-020-00318-5.
[30] Z. Khokhar and M. A. Siddiqi, “Machine Learning Based Indoor Localisation Using Wi-Fi and Smartphone,” *Journal of Independent Studies and Research Computing*, 2020, doi: 10.31645/06.
[31] M. Hasan, M. M. Islam, M. I. I. Zarif, and M. M. A. Hashem, “Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches,” *Internet of Things*, vol. 7, p. 100059, Sep. 2019, doi: 10.1016/j.iot.2019.100059.
[32] G. Nicora, M. Rios, A. A.-Hanna, and R. Bellazzi, “Evaluating pointwise reliability of machine learning prediction,” *Journal of Biomedical Informatics*, vol. 127, p. 103996, Mar. 2022, doi: 10.1016/j.jbi.2022.103996.
[33] L. Antony *et al.*, “A Comprehensive Unsupervised Framework for Chronic Kidney Disease Prediction,” *IEEE Access*, vol. 9, pp. 126481–126501, 2021, doi: 10.1109/access.2021.3109168.
[34] N. Rajagopal, “Localization, Beacon Placement and Mapping for Range-based Indoor Localization Systems - ProQuest,” *Pro Quest LLC*, Aug. 2019, [Online]. Available: https://www.proquest.com/openview/5ecde497dafe1da4cf958f3dae857772/1?pq-origsite=gscholar&cbl=18750&diss=y.
[35] R. Chen *et al.*, “Precise Indoor Positioning Based on Acoustic Ranging in Smartphone,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021, doi: 10.1109/tim.2021.3082269.
[36] B. Jang, H. Kim, and J. W. Kim, “Survey of Landmark-based Indoor Positioning Technologies,” *Information Fusion*, vol. 89, pp. 166–188, Jan. 2023, doi: 10.1016/j.inffus.2022.08.013.

BIOGRAPHIES OF AUTHORS

Kamel Maaloul obtained an engineer in computer media from Skikda University, Skikda, Algeria in 2004, and a master's in artificial intelligence from El-Oued University, El-Oued, Algeria in 2017. He is currently a PhD student from El-Oued University, El-Oued, Algeria. And a principal engineer in computer media at El-Oued University, El-Oued, Algeria. He can be contacted at email: maaloul-kamel@univ-eloued.dz.

Brahim Lejdel is a full Professor in the Faculty of Exact Sciences, University of EL-Oued (Algeria). He has held a Magister Degree in Computer Science from University of Ouargla since 2009. He held his PhD in computer science in 2015. He is the author of more than 120 published articles in refereed journals and International Conferences. He publishes four books in his research domain. He is the chair of many international conferences. He works as an invited editor, reviewer, and others in many journals and conferences. He is the supervisor of many students in computer science in license, master and doctorate. He can be contacted at email: lejdel.brahim@gmail.com.

Eliseo Clementini is an associate professor of computer science at the Department of Industrial and Information Engineering and Economics of the University of L'Aquila (Italy). He received a M.Eng. in Electronics Engineering from University of L'Aquila in 1990 and a Ph.D. in Computer Science from University of Lyon (France) in 2009. He has been a visiting professor at the National Center for Geographic Information and Analysis of the University of Maine, at the Department of Geography and Geomatics of the University of Glasgow, at the Department of Geography of the University of Liege, and at the Department of Geodesy and Geoinformation of Vienna University of Technology. His research interests are mainly in the fields of spatial databases and geographical information science. He contributed to open geospatial consortium (OGC) recommendations and ISO/TC 211 standard. He has been an invited speaker in various international symposia. He is in the editorial board of the ISPRS International Journal of Geo-Information and is a member of the program committee of main international GIS conferences. He can be contacted at email: eliseo.clementini@univaq.it.

Nedioui Med Abdelhamid is an associate Professor in the Faculty of Exact Sciences, University of EL-Oued (Algeria). He has held a Magister Degree in Computer Science from University of Biskra since 2015. He held his PhD in computer science in 2021. He can be contacted at email: nedioui3904@gmail.com.