Comparative Analysis of IoT Enabled Multi Scanning Parking Model for Prediction of Available Parking Space with Existing Models

Anchal[1†] and Pooja Mittal[2††],
aanchaldahiyaa@gmail.com mpoojamdu@gmail.com
Maharshi Dayanand University, Department of Computer Science, Rohtak, India

Abstract
The development in the field of the internet of things (IoT) have improved the quality of the life and also strengthened different areas in the society. All cities across the world are seeking to become smarter. The creation of a smart parking system is the essential use case in smart cities. In recent couple of years, the number of vehicles has increased significantly. As a result, it is critical to make the use of technology that enables hassle-free parking in both public and private spaces. In conventional parking systems, drivers are not able to find free parking space. Conventional systems requires more human interference in a parking lots. To manage these circumstances there is an intense need of IoT enabled parking solution that includes the well defined architecture that will contain the following components such as smart sensors, communication agreement and software solution. For implementing such a smart parking system in this paper we proposed a design of smart parking system and also compare it with convetional system. The proposed design utilizes sensors based on IoT and Data Mining techniques to handle real time management of the parking system. IoT enabled smart parking solution minimizes the human interference and also saves energy, money and time.
Keywords:
IoT, IoT Enabled Smart Parking, Data Mining, Detection agent, Smart Parking Agent.

1. Introduction

As the Indian population grows and people's lifestyles change, every household now owns a vehicle. The rapid growth of India's population has resulted in a quick rise of the vehicles on the road. Inconvenience in public transit is also a big factor in the rise in the number of vehicles on the road. Due to a growing population, parking space in Indian cities has stayed stable or decreased. There is currently no standard technique for finding a parking place in a parking lot. Existing technologies are slow and relying on human interaction. Traffic congestion is caused by people looking for a parking spot.

The work is processed towards the fulfilment of the proposed framework for our model. The complete system is design to consider two different types of agents which reacts as per the learning and retrieval results. One is the detection agent, and the second is the smart parking agent. The working of the detection agent is to identify the type of vehicle. On the other hand, the smart parking agent's work is to allot the free parking space to the specific vehicle type according to the vehicle id, which is generated by the detection agent. The objective behind the smart parking system is to utilize the free parking space effectively. This method helps us to efficiently arrange the available parking space, especially in the peak areas.

1.1 Organization

The paper is arranged in the following way. In section 2 literature survey about the work is highlighted. In section 3 IoT technologies are introduced. In section 4 problem is formalized. In section 5 we will discuss about the research methodology and also introduce the multi scanning model. In section 6 we will discuss about the dataset that is used for analysis. In section 7 results are evaluated on the basis of 4 parameters that are accuracy, F-1 score, recall and precision and comparison is performed with previous models and in the last section 8 conclusion is presented.

2. Literature Survey

Takizawa et al. [1] proposed a smart parking system that uses closed-circuit television to access the place of a vehicle in a parking lot. Pixel recognition is used to decide the location. Setting threshold values for pixels identifies the difference between the absence and presence of vehicles in the parking slot, and images are considered as grayscale images. This method is more accurate, reliable, and consistent, and it can be used to collect information about available parking spaces.

Bong et al. [2] proposed an image processing method called Car Occupancy Information System (COINS) to find out if the car is parked or if the parking space is free. The benefit of using this approach is the COINS are cheap because the sensors have not been used. At the same time, using the procedures of image processing lowers the system's cost. The magnetic field of the Earth is used to locate empty parking spaces. To obtain the necessary

Manuscript received August 5, 2022
Manuscript revised August 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.8.51

information, the computers that control the structure are linked with different sensors.
Geng and Cassandras [3] implemented the smart parking system in the urban areas by considering the environmental factors. The driver's cost feature is determined based on how near the people is to the destination and the cost of the parking area. The cost feature is used to assign and reserve the best parking spot for the people who have made the request. The problem of Mixed Integer Linear Programming (MILP) is overcome using this method at any point in time when a decision must be made. The elucidation of MILP is the most appropriate allocation of the parking area based on the data available at the time. This detail is brought up during the next instance's decision-making process. It is ensured that conflicts in resource reservation are avoided, as well as that the parking slot is only given to the user with the highest cost function. It has been established that the time it takes to locate a parking space is reduced, as is the charge for that space. Simultaneously, it is demonstrated that the technique is efficient in efficiently using the parking room.

Pham et al. [4] proposed an algorithm that improves the efficiency of the current intelligent parking system based on the cloud approach. IoT is often used to build the system's architecture. The algorithm will quickly identify vacant parking spaces and is fully automated. The cost of parking for users is determined by the distance and amount of available spaces in the parking area. The user may be assigned a parking spot based on the cost if they request it. If the requested parking area is complete, the user can be directed to another parking area. Furthermore, the author proposed an algorithm in order to decreases the user's wait time and increases the chances of finding a parking spot.
Sarangi et al. [5] proposed an IoT-based smart parking system. The system is designed to identify the vehicle that arrived at the parking slot, in order to verify the vacant space in the parking areas and to provide an interface to the user. The system consists of three different approaches: computer vision procedures, wireless sensor networks, and android mobile. Users will use the Android-based application to locate empty parking spaces and pay for them. In the construction of smart cities, Kumar et al. [6] used a combination of two technologies called radar and ultrasonic detection to introduce a smart parking system. The main goal of the model is to create a self-governing parking system that reduces traffic jams, air pollution, and user waiting time by allowing them to easily find parking slots even during peak hours. In this new parking system solution, the human presence would be minimized. The Wi-Fi system is used to transfer all of the data for storage to the cloud. With the help of the interface, the user can find an empty parking spot. The device proposed by the author has continuously monitored the vacancy and occupancy of each parking slot. When the vehicles in a parking location exceed the amount of parking locations in the parking area, the device will sound an alarm. The same device can be used to provide a cost-effective and efficient parking procedure in colleges, shopping centres, movie theatres, and other locations.

3. IoT Technologies

The term IoT was presented by Kevin Asthon in 1999 [17]. As technology progresses, it has the potential to connect everything that is connected by networks and facilitate communication with less human intervention. There is no common design available in current state because IoT is in the beginning stage. The IoT has several definitions due to non-existing guidelines for IoT definition. It is described as the physical environment that is associated to the web via wireless or wired networks and integrated procedures or sensors. The associated devices are called smart devices. IoT is concerned with connecting the entire globe via the Internet. IoT allows billions of nodes from various things to be connected to the big superstore web browsers and hubs. IoT also aids in the integration of new software and communication networks. The primary objective of the Internet of Things is to make the world smarter by providing the data it requires through ancient and real-time feeds, using data mining algorithms to provide the data it requires, and instantaneously implementing computational expertise intelligence to make intelligent decisions.

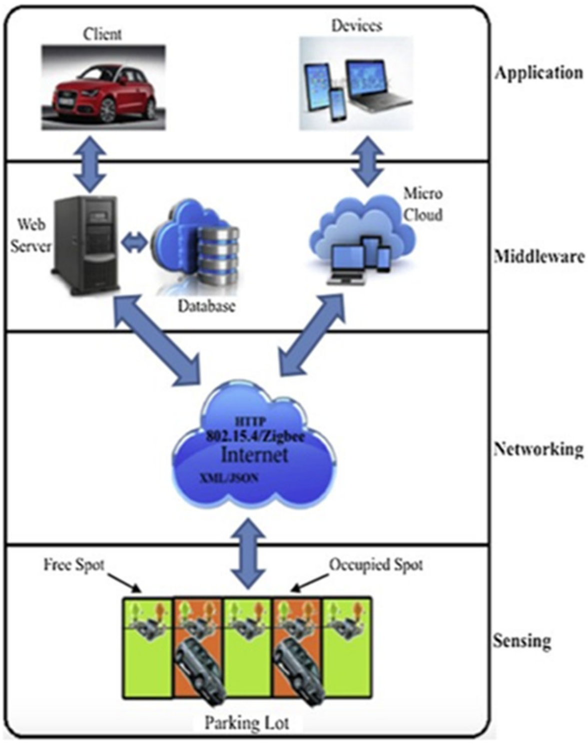

Fig. 1. IoT Enabled Parking Layers [13]

4. Problem Formulation

To find a free location to park a vehicle during peak hours is become one of the difficult task. . To park the vehicle whether it is in shopping malls, restaurants, or offices, etc., is a challenging process and also leads to the wastage of valuable resources. Many works are promoted as smart parking systems in the literature, however they fail to address the major obstacles. User privacy, easy identification of parking places, presentation of an optimal and real-time significant system, and other difficulties will be highlighted. The majority of the issues raised are being considered for a solution in the current work. The current study provides a prediction system that use data mining techniques to provide quick and easy access to information. The work also considers and presents a predictive process on the basic part history records of the vehicle to predict the parking slot. A supervised approach is examined for entire prediction results. The entire system was created with two sorts of agents in mind: a "detection agent" and a "smart parking agent", both of which react based on learning and retrieval outcomes.

5. Research Methodology

Figure 2 depicts a block diagram of the suggested multi scanning technology. We proposed a multi scanning strategy since sensors create a vast quantity of data in every second and we require to collect the data that is relevant to parking. The main goal of the smart parking system is to maximize the use of free parking space. The methodology is broken down into two sections:

- Detection agent
- Smart parking agent

Vehicle type is identified by the detection agent. User submits a request for a parking spot that is available. Following receiving of the user's request, the system begins to determine the type of vehicle, such as sedan, coupe, SUV, bus, auto, and so on, and assigns a unique vehicle id to each vehicle. If a slot in the parking spot is available, the smart parking agent assigns a specific location to the vehicle in the next section. The smart parking agent classifies the spaces in the parking areas in advance based on the kind of vehicle. Smart parking agent will assign the available parking slot to the vehicle according to the classification of slot. Classifiers used by smart parking agent are SVM, Decision tree, Random forest etc.

6. Data Set

Primary step is to collect the data from the sensors that are located in various parking areas in the smart cities. In our work data is collected as a CSV file which is from Birmingham city in United Kingdom in various car park slots. Data is collected over 3 month of data. The main objective is the prediction of available space within the time interval of 10minute to 20 minute. The data is organized in the following way.

1. Parking unique Id: Unique id is assigned to each parking slot.
2. Time: Timestamp during which data is collected.
3. Duration: Total amount of time for which a slot is engaged or free.
4. Start Time and Finish Time: Time at which slot is assigned and time at which slot is free.
5. Status: Parking area is free or allotted is updated using status queue.

Table 1 Input Features

Key Features	*Range*
Unique ID	Unique id assigned to sensors in every slot
Day	Number of days in weak (1-7) days.
Hours(Start)	24 hours of a day
Hours(Finish)	24 hours of a day
Minutes(Start)	0-59 min. of an hour
Minutes(Finish)	0-59 min. of an hour
Status	Slot is free or engaged

7. Result Evaluation

In this paper we will evaluate 3 data mining techniques. Results are evaluated on the basis of the following techniques.

7.1 Decision tree

Decision tree is made up of several layers of testing nodes, branches that signify the result of testing node, and child nodes. In classification problems, the child nodes reflect the class label [13].

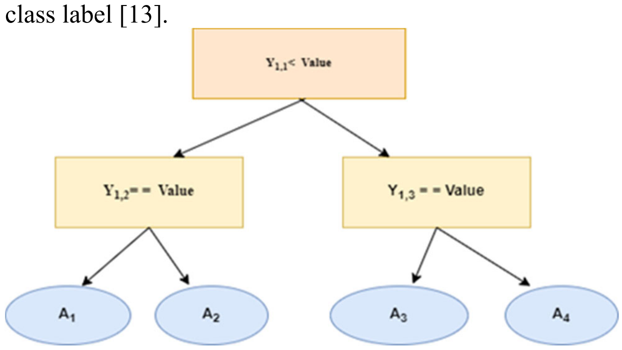

Figure 3 Basic working of decision tree

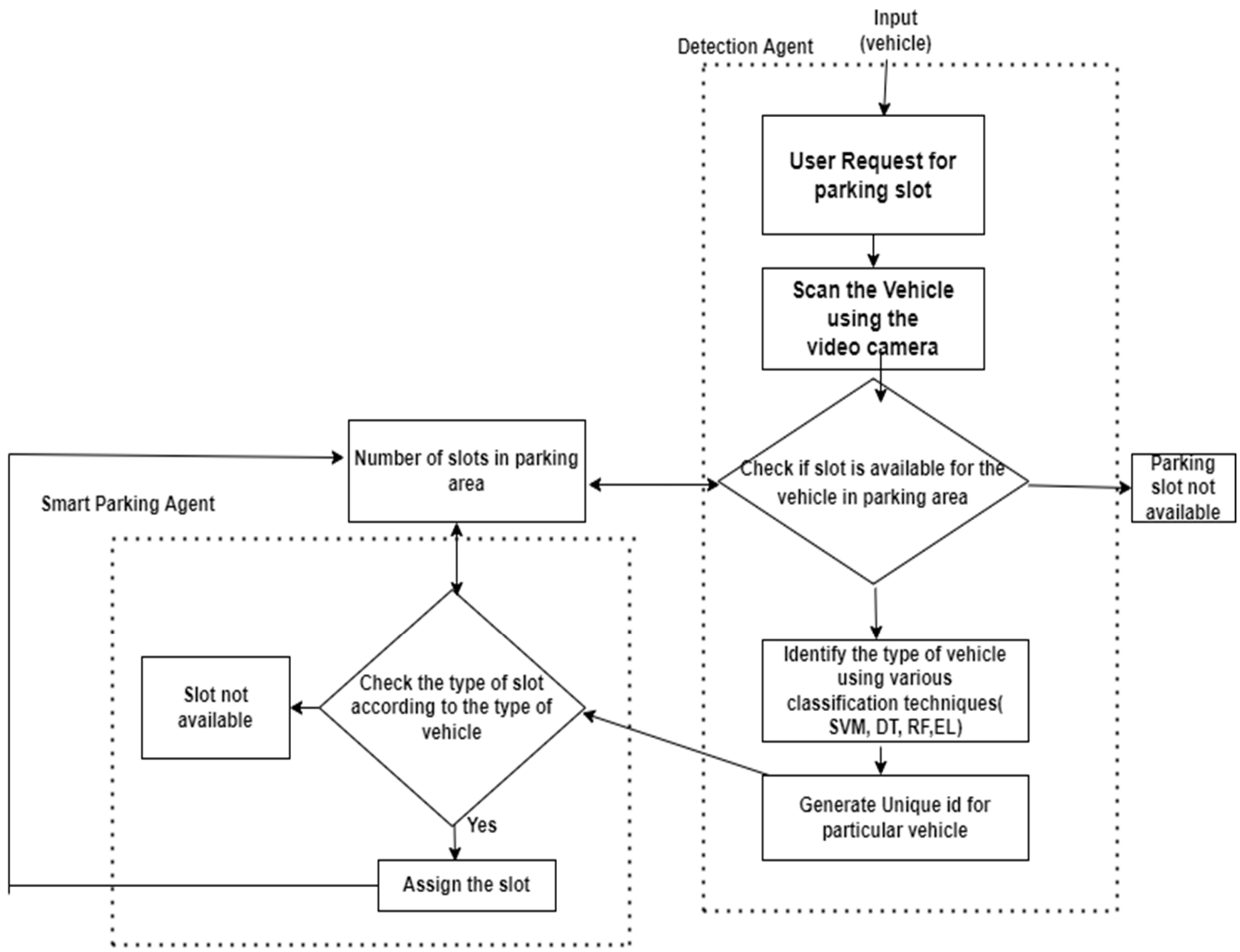

Figure 2. IoT Enabled Multi Scanning Parking Model

The rules are represented by the path from root to leaf. Predictive models based on trees provide great accuracy, consistency, and applicability. Basic architecture of decision tree is shown in the figure 3.

7.2 Logistic Regression

For predictive analytics and modelling, a logistic model is often adopted. In this analytics technique, the dependent variable is either finite or categorical. In statistical software, logistic regression is used to determine the association among a dependent variable and one or more independent variables. It is supervised learning algorithm. Also known as discriminative forecasting model. To represents the subsequent likelihood dissemination P(Y|X), where Y represents the target variable and X represents the characteristics. Here X, returns the likelihood dissemination over the value. The given figure shows the architecture of logistic regression.“The outcome of the sigmoid function in a binary classification method is read as the likelihood of a specific model that belongs to a positive class, i.e., f(z) = P(y = 1jx;w), where z is the linear combination of weights and sampling features z = wTx”.

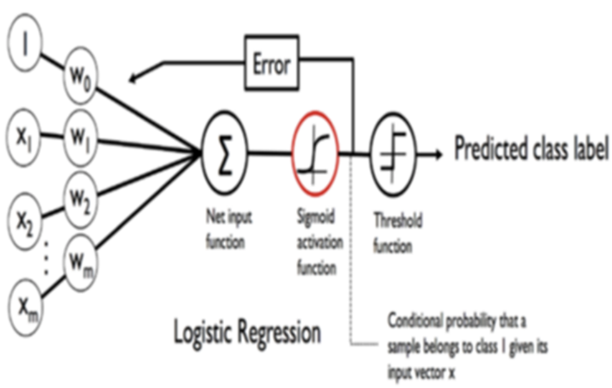

Fig 4 Logistic Regression Basic architecture

7.3 Support Vector Machine

It's a supervised learning technique that uses learning algorithms to evaluate data for classification and regression analysis. It's one of the empirical learning-based predictive models out there. It can take two groups of data that is training data and testing data. To optimize the degree of a space between these two groups, the SVM learning algorithm transfers the training data sets samples to points in space [13]. The additional values are now plotted into the similar space, and their group membership is projected depending on which side of the gaps they lie on. Purpose of SNM algorithm is to find the optimal line or decision boundary for classifying n-dimensional space into classes so that supplementary data points can be voluntarily placed in the accurate group in the long term. A hyper plane is the name for the ideal choice limit. Pictorial representation of workflow of SVM is shown in figure 5." In the figure the SVM takes the data from the database that will contain heterogeneous data, it maps the training data into different groups according to their category".

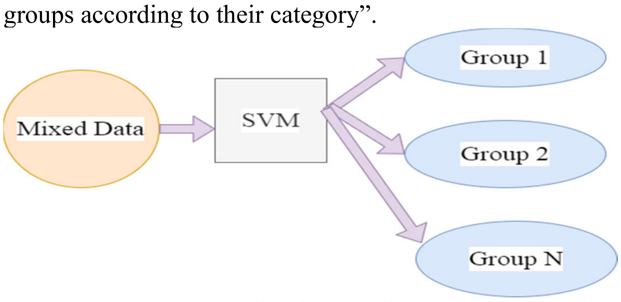

Fig. 5 Block diagram of SVM

7.4 Hyper Parameters of Data Mining Techniques

Table 2 represents the hyper parameters of the 3 data mining techniques that are tuned for comparative analysis. Grid search [14] is used for best result of hyper parameters values for each data mining technique. Hyper parameters are the internal coefficients for a model found by a learning algorithm.

Table 1 Tuning of Hyper Parameters of Data Mining Techniques

Logistic Regression		***Decision Tree***		***SVM***	
Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
Solver	Lib linear	Max depth	10	Kernal	linear
Penalty	12	Min-sample splits	(2,30, 2)	Degree	3
C	100	Criterion	Gini Entropy	C	prange (0.001, 1000.0, 10) **100.0**

We have tuned by 3 parameters in Logistic Regression. The parameters that are used for tuning are Solver whose value is set to lib linear. Another parameter used is penalty ard it is set by 12 initially. And the last parameter used for tuning internal coefficient of logistic regression is C and it is initialize by 100. Similarly for Decision tree we tuned 3 hyper parameters. Max-depth states the maximum deepness of the tree and min-sample-leaf defines number of samples contained by the leaf node. Criterion= entropy it works on information gain that belong to the number of splits in the decision tree. Similarly SVM is tuned by 3 parameters here kernel is set to linear we can also set it to rbf according to our need. And the degree is set to 3. High priority is assigned to SVM because the performance of SVM is better among all the data mining techniques.

7.5 Parameters used for Evaluation

The parameters that we used for evaluation and comparing the multi scanning data mining model are given below. For checking the stability and over fitting problem of the model we used k-fold cross validation method.

Accuracy: Most common used metrics to judge any model is known as accuracy metric. It measures the accurately forecasted samples among all the given samples. The following equation is used.

$$\text{Accuracy} = \frac{\text{Accurate Prediction}}{\text{Total Samples}} \tag{1}$$

F1- Score: It is harmonic mean of precision and recall [15]. It uses both the metrics. Higher the F1 score the better the results are.

$$\text{F1} - \text{Score} = \frac{2*(\text{Recall}*\text{Precision})}{\text{Recall}+\text{Precision}} \tag{2}$$

Recall: It is defined as a ratio of positive instance out of all the total actual positive instances. The following equation is used.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive}+\text{False Negative}} \tag{3}$$

Precision: It is defined as a ratio of positive instances out of total forecasted positive instances.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Postive} \quad (4)$$

K-fold-cross-validation: It is a technique for detecting over fitting data and determining the accuracy of the model. K-fold means, data is divided into K equal sets. Among all the given K sets each set is processed only once as testing data, with the subsequent sets being used as training data. 5-fold cross-validation is used in this paper.

7.6 Performance Evaluation

In this section we will evaluate the performance of Logistic Regression, Decision tree and Support Vector Machine (SVM) algorithms in terms of result related to each cross validation. A comparison of 10-minute and 20-minute prediction was conducted, with 65 percent and 85 percent threshold used for each.

7.6.1 10 Minute Prediction (65% Threshold)

In Table 3 represents the average cross- validation score of Logistic Regression, Decision tree and SVM models given 10-minute predictions with 65% threshold. Given table clearly shows that the performance of logistic regression is least during comparative analysis. The average performance is 70.56% average accuracy, average F1-Score is 55.35%, average recall is 60.445% and average precision is 75.45%. The performance of Decision tree and support vector machine is quite close to each other with average accuracy 93.80%, average F1_score is 91.80%, average recall is 90.23%, and average precision is 93.27%. Support vector machine outperform among all the techniques with average accuracy 93.75%, average F1-score 91.70%, average recall is 91.30 % and average precision is 92.14%.

Table3: Average Cross Validation Result of Each Model

Metrics	***"Logistic Regression"***	***"Decision Tree"***	***"Support Vector Machine"***
Accuracy	70.56	93.80	93.75
F1- Score	55.35	91.80	91.70
Recall	60.45	90.23	91.30
Precision	75.45	93.27	92.14

7.6.2. 10 Minute prediction (85% Threshold)

In the table 4 represents the average cross-validation score of Logistic Regression, Decision Tree and SVM models given 10- minute prediction with 85% threshold. By following 65% threshold the performance of logistic regression is poor among all the techniques. Average accuracy of logistic regression is 70.89%, average F1-Score is 55.05%, average recall value is 60.20% and average precision is 75.57%. The performance of decision tree average accuracy is 94.09%, average F1-score is 91.50%, average recall is 89.12% and average precision is 94.03%. The performance of Support vector machine is best among all the technique here average accuracy is 94.25%, average F1-score is 90.60%, average recall value is 91.20%, and average precision is 92.10%.

Table 4 Average Cross Validation Result of Each Model

Metrics	***"Logistic Regression"***	***"Decision Tree"***	***"Support Vector Machine"***
Accuracy	70.89	94.09	94.25
F1- Score	55.05	91.50	90.60
Recall	60.20	89.12	91.20
Precision	75.57	94.03	92.10

7.6.3. 20 Minute Prediction (65% Threshold)

In this section we will represent a comparative analysis using 20 minute prediction having 65% threshold. In table 5 the average accuracy of logistic regression is 73.56%, average F1-score is 60.35%, average recall is 65.45% and average precision is 76.16%. The performance of Decision tree and support vector machine is quite close to each other with average accuracy 89.12%, average F1_score is 87.05%, average recall is 84.12%, and average precision is 89.10%. Support vector machine outperform among all the techniques with average accuracy 88.24%, average F1-score 86.01%, average recall is 85.10 % and average precision is 86.14%.

Table 5 Average Cross Validation Result of Each Model

Metrics	*"Logistic Regression"*	*"Decision Tree"*	*"Support Vector Machine"*
Accuracy	73.56	89.12	88.24
F1- Score	60.35	87.05	86.01
Recall	65.45	84.12	85.10

Precision	76.16	89.10	86.14

7.6.4. 20 Minute Prediction (85% Threshold)

In the table 6 represents the average cross-validation score of Logistic Regression, Decision Tree and SVM models given 20- minute prediction with 85% threshold. By following 65% threshold the performance of logistic regression is poor among all the techniques. Average accuracy of logistic regression is 74.10%, average F1-Score is 60.26%, average recall value is 65.10% and average precision is 77.75%. The performance of decision tree average accuracy is 89.20%, average F1-score is 86.10%, average recall is 83.19% and average precision is 90.05%. The performance of Support vector machine is best among all the technique here average accuracy is 88.60%, average F1-score is 85.04%, average recall value is 85.15%, and average precision is 86.24%.

Table 6 Average Cross Validation Result of Each Model

Metrics	*"Logistic Regression"*	*"Decision Tree"*	*"Support Vector Machine"*
Accuracy	74.10	89.20	88.60
F1- Score	60.26	86.10	85.04
Recall	65.10	83.19	85.15
Precision	77.75	90.05	86.24

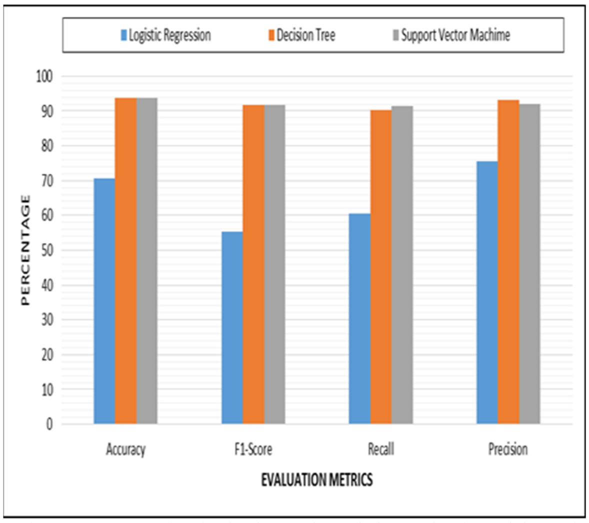

Fig. 6 Representation in the form of graph for evaluation of data mining techniques using 65% threshold.

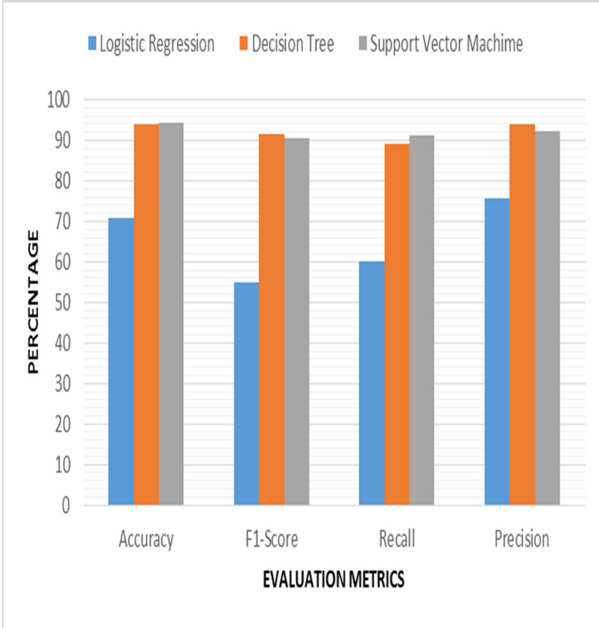

Fig. 7 Representation in the form of graph for evaluation of data mining techniques using 85% threshold.

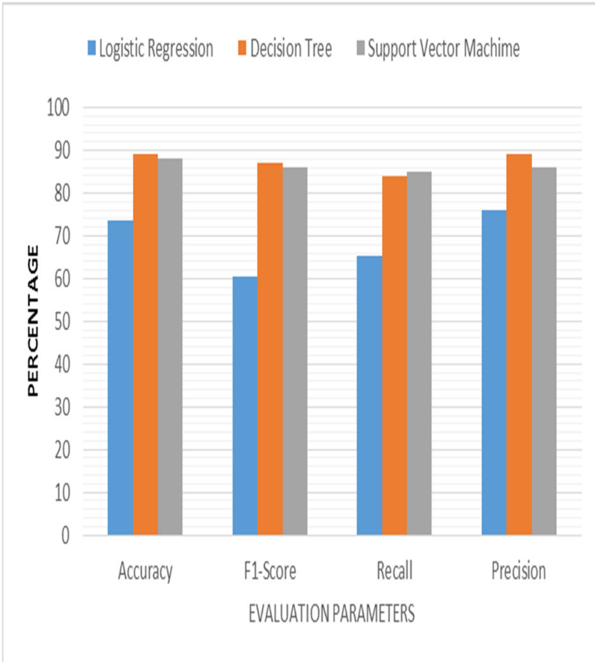

Fig. 8 Representation in the form of graph for evaluation of data mining techniques using 65% threshold.

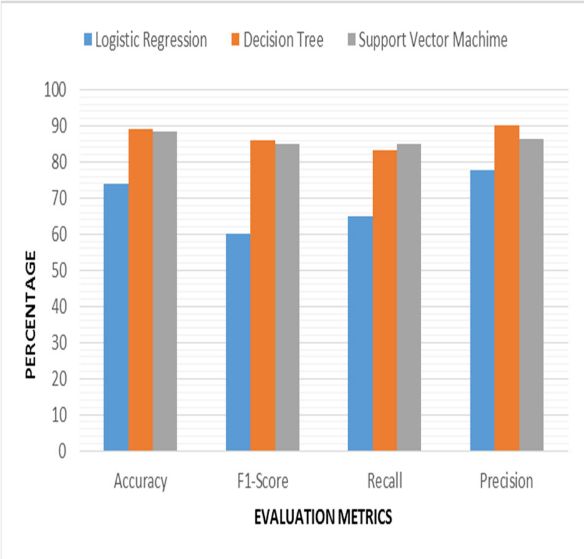

Fig. 9 Representation in the form of graph for evaluation of data mining techniques using 85% threshold.

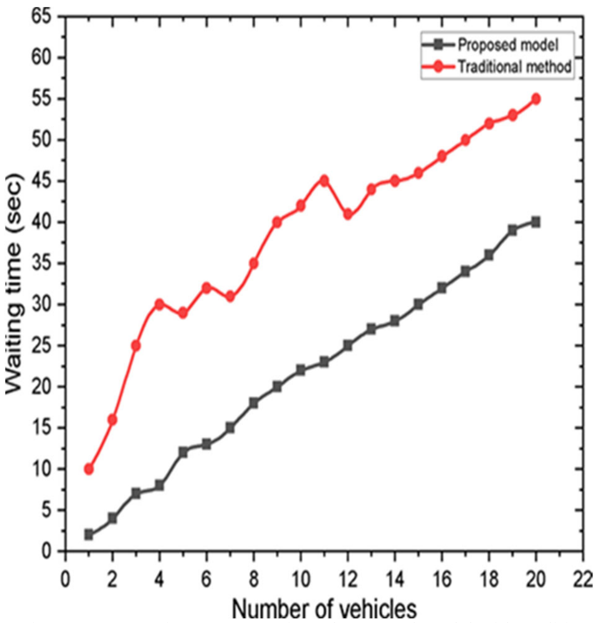

Fig. 10 Comparative representation of proposed model with traditional model

Figure 10 shows the comparative representation how our model saves time as compared to traditional method. The measures that are selected for comparison are area it is used to help in finding the free parking space. Our proposed model use sensor data that will detect the vehicle in advance. It also facilitates vehicle counting system that will detect the number of free space according to the type of vehicle.
Tables are used to organize the following features which are used as input features for our data mining model.

7.6.5 Comparison of Models

Table 7 Represents Comparison with Conventional Model

S. No	*Key Issues*	*Proposed Model*	*Conventional Models*
1	Area	10	10
2	Type of Sensor Used	Ultrasonic Sensors	Arduino Uno
3	Vehicle Count System	Yes	No
4	Car Parking Management	Yes	Yes
5	Technology Used	Yes	No
6	Real Time Parking Tracking	Yes	No
7	Optimization of Space	Yes	No
8	Congestion	Congestion can be decreased by diverting traffic to available spaces.	Congestion increase due to parking traffic.
9	Scanning	Sensors emit a significant amount of data, we use multi scan to collect only the data that is relevant for real-time parking.	No Scanning is used

8. Conclusion

This paper shows the IoT based smart parking model to help the drivers to find the free parking space smartly without wasting precious time and also prevent from human intervention. Basically the proposed model works in two steps that is detection agent and smart parking agent. Detection agent helps in identification of the type of the vehicle and smart parking agent helps the drivers to assign the parking slot if slot is available according to the type of the vehicle detected by the detection agent. Smart parking agent uses data mining techniques to predict the available parking slot. In the model, different technologies are converged that includes commodity sensors, real-time analytics, and smart identification of type of vehicle. Internet of things (IoT) is used that mechanize existing parking system in the real world. A comparative analysis is also done that will distinguish the model with traditional model. Four parameters are used for evaluation of the result. These parameters are Accuracy, F1- Score, Recall and Precision.

References

[1] Takizawa, H., Yamada, K., & Ito, T. (2004, June). Vehicles detection using sensor fusion. In IEEE Intelligent Vehicles Symposium, 2004 (pp. 238-243). IEEE.
[2] Bong, D. B. L., Ting, K. C., & Lai, K. C. (2008). Integrated Approach in the Design of Car Park Occupancy Information System (COINS). IAENG International Journal of Computer Science, 35(1).
[3] Geng, Y., & Cassandras, C. G. (2013). New "smart parking" system based on resource allocation and reservations. IEEE Transactions on intelligent transportation systems, 14(3), 1129-1139.
[4] Pham, T. N., Tsai, M. F., Nguyen, D. B., Dow, C. R., & Deng, D. J. (2015). A cloud-based smart-parking system based on Internet-of-Things technologies. IEEE Access, 3, 1581-1591.
[5] Sarangi, M., Mohapatra, S., Tirunagiri, S. V., Das, S. K., & Babu, K. S. (2020). IoT aware automatic smart parking system for smart city. In Cognitive Informatics and Soft Computing (pp. 469-481). Springer, Singapore.
[6] Kumar, I., Manuja, P., Soni, Y., & Yadav, N. S. (2020). An Integrated Approach towards Smart Parking Implementation for Smart Cities in India. In Advances in Data and Information Sciences (pp. 343-349). Springer, Singapore.
[7] Umbaugh, S. E. (1997). Computer vision and image processing: a practical approach using cviptools with cdrom. Prentice Hall PTR.
[8] Yang, F., & Ma, Z. (2005, October). Vehicle license plate location based on histogramming and mathematical morphology. In Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05) (pp. 89-94). IEEE.
[9] Mehena, J. Vehicle License Plate Segmentation and Extraction Based on Mathematical Morphology.
[10] Lumonics, G. S. I. (1999). QuantArray analysis software. Operator's Manual.
[11] Goyal, A., & Verma, B. (2007, April). A neural network based approach for the vehicle classification. In 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (pp. 226-231). IEEE.
[12] Huang, K.; Chen, K.; Huang, M.; Shen, L. Multilayer perceptron with particle swarm optimization for well log data inversion. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 6103–6106. .
[13] Anchal D., Pooja M. (2021) Evaluation of Data Mining Techniques and Its Fusion with IoT Enabled Smart Technologies foe Effective Prediction of Available Parking Space (pp. 187-197) , IJECES.
[14] Goyal, A., & Verma, B. (2007, April). A neural network based approach for the vehicle classification. In 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (pp. 226-231). IEEE
[15] Lipton, Z.C.; Elkan, C.; Naryanaswamy, B. Optimal thresholding of classifiers to maximize F1 measure. In Proceedings of the Joint European Conf. on Machine Learning and Knowledge Discovery in Databases, Nancy, France, 15–19 September 2014; pp. 225–239.

Ms. Anchal has completed her MCA from MD University . She is currently pursing Ph.D in Computer Science at Department of computer Science and Applications M.D. University, Rohtak. Her main research area includes Internet of things and Data Mining.

Dr. Pooja Mittal is currently working as Assistant Professor at Department of computer Science and Applications, M.D. University, Rohtak, India. She obtained her Ph.D from Maharshi Dayanand University. Her area of research and specialization includes Data Mining, Data Warehousing and Computer Science. She has published more than 50 research papers in renowned international/ national journals and attended more than 40 conferences.